

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2005, 08:40:57 ; Search time 199.724 Seconds  
(without alignments)  
1690.542 Million cell updates/sec

Title: US-10-624-932-2\_COPY\_26\_898  
Perfect score: 4660  
Sequence: 1 QQSATVANPVPGANPDLLPH.....AVAGLGQPDAGLFTVSEAEC 873

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%						
		Query						
Result	No.	Score	Match	Length	DB	ID	Description	
	1	4660	100.0	898	5	AAU85403	Aau85403 Human pro	
	2	4660	100.0	898	8	ADH71618	Adh71618 Human pro	
	3	4650	99.8	898	5	AAU97899	Aau97899 Human net	
	4	4567.5	98.0	899	5	AAU79939	Aau79939 Human UNC	
	5	4567.5	98.0	899	7	ADG42569	Adg42569 Novel hum	
	6	4567.5	98.0	899	8	ADH71610	Adh71610 Human pro	
	7	4567.5	98.0	899	8	ADH71628	Adh71628 Human pro	
	8	4566.5	98.0	899	8	ADH71636	Adh71636 Human pro	

9	4564.5	98.0	899	8	ADH71642	Adh71642	Human	pro
10	4564.5	98.0	899	8	ADH71640	Adh71640	Human	pro
11	4563.5	97.9	899	8	ADH71632	Adh71632	Human	pro
12	4563.5	97.9	899	8	ADH71638	Adh71638	Human	pro
13	4562.5	97.9	899	8	ADH71634	Adh71634	Human	pro
14	4560.5	97.9	899	8	ADH71648	Adh71648	Human	pro
15	4560.5	97.9	899	8	ADH71630	Adh71630	Human	pro
16	4560.5	97.9	899	8	ADH71644	Adh71644	Human	pro
17	4555.5	97.8	899	8	ADH71646	Adh71646	Human	pro
18	4553.5	97.7	899	8	ADH71650	Adh71650	Human	pro
19	4552	97.7	898	8	ADH71626	Adh71626	Human	pro
20	4514.5	96.9	943	4	AAM79128	Aam79128	Human	pro
21	4511	96.8	898	2	AAW78898	Aaw78898	Rat	UNC-5
22	4511	96.8	898	5	AAU10543	Aau10543	Rat	netri
23	4511	96.8	898	5	AAU97900	Aau97900	Rat	netri
24	4493	96.4	898	7	ADG42580	Adg42580	Rat	trans
25	4282	91.9	833	8	ADH71622	Adh71622	Human	pro
26	4282	91.9	842	5	AAU74818	Aau74818	Human	REP
27	4282	91.9	842	8	ADL06574	Adl06574	Human	tum
28	4275	91.7	817	8	ADH71624	Adh71624	Human	pro
29	2839	60.9	544	7	ADG42581	Adg42581	Human	tra
30	2815	60.4	556	2	AAW78899	Aaw78899	Human	UNC
31	2759.5	59.2	931	7	ADG42582	Adg42582	Mouse	tra
32	2747.5	59.0	931	4	AAB50691	Aab50691	Human	UNC
33	2747.5	59.0	931	7	ADE63098	Ade63098	Human	Pro
34	2747.5	59.0	931	7	ABU64297	Abu64297	Human	thr
35	2747.5	59.0	931	8	ADR99258	Adr99258	Human	unc
36	2747.5	59.0	982	4	ABG11551	Abg11551	Novel	hum
37	2736	58.7	964	8	ADR99250	Adr99250	Human	lRO
38	2733.5	58.7	931	7	ADG42584	Adg42584	Human	tra
39	2688.5	57.7	929	7	ADG42583	Adg42583	Human	tra
40	2560.5	54.9	943	2	AAW78900	Aaw78900	Rat	UNC-5
41	2560.5	54.9	945	7	ADE63096	Ade63096	Rat	Prote
42	2546.5	54.6	933	5	AAO18734	Aao18734	Human	NOV
43	2546.5	54.6	933	5	AAO18735	Aao18735	Human	NOV
44	2546.5	54.6	945	5	ABB09520	Abb09520	Human	tra
45	2541.5	54.5	945	4	AAU12244	Aau12244	Human	PRO

# ALIGNMENTS

## RESULT 1

AAU85403

ID AAU85403 standard; protein; 898 AA.

XX

AC AAU85403;

XX

DT 21-MAY-2002 (first entry)

XX

DE Human protein NOV1.

XX

KW Human; NOVX; cardiomyopathy; atherosclerosis; diabetes;

KW cell signal processing disorder; metabolic disorder; obesity; infection;

KW anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;

KW Alzheimer's disease; Parkinson's disease; immune disorder;

KW haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;

KW osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;  
 KW myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;  
 KW psychosis; neurological disorder; anxiety; schizophrenia;  
 KW manic depression; dementia; dyskinesia; Huntington's disease;  
 KW Gilles de la Tourette's syndrome; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200210216-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 30-JUL-2001; 2001WO-US024225.  
 XX  
 PR 28-JUL-2000; 2000US-0221409P.  
 PR 04-AUG-2000; 2000US-0222840P.  
 PR 04-AUG-2000; 2000US-0223752P.  
 PR 04-AUG-2000; 2000US-0223762P.  
 PR 04-AUG-2000; 2000US-0223769P.  
 PR 04-AUG-2000; 2000US-0223770P.  
 PR 14-AUG-2000; 2000US-0225146P.  
 PR 15-AUG-2000; 2000US-0225392P.  
 PR 15-AUG-2000; 2000US-0225470P.  
 PR 16-AUG-2000; 2000US-0225697P.  
 PR 01-FEB-2001; 2001US-0263662P.  
 PR 05-APR-2001; 2001US-0281645P.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Padigar M, Mezes P, Mishra V, Burgess C, Casman S, Grosse WM;  
 PI Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;  
 XX  
 DR WPI; 2002-180074/23.  
 DR N-PSDB; ABK37922.  
 XX  
 PT New isolated cytoplasmic, nuclear, membrane bound, or secreted  
 PT polypeptide, useful for treating cardiomyopathy, atherosclerosis,  
 PT infections, cancer, neurodegenerative, metabolic, hematopoietic and  
 PT immune disorders.  
 XX  
 PS Claim 1; Page 11; 213pp; English.  
 XX  
 CC The invention relates to an isolated cytoplasmic, nuclear, membrane  
 CC bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature  
 CC form. Also included are the nucleic acids encoding the NOVX proteins, a  
 CC vector comprising the nucleic acid, a cell comprising the vector, an anti  
 CC -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the  
 CC antibody are useful for treating or preventing a NOVX-associated  
 CC disorder, where the disorder is selected from cardiomyopathy,  
 CC atherosclerosis, diabetes, a disorder related to cell signal processing  
 CC and metabolic pathway modulation, metabolic disorders, obesity,  
 CC infectious disease, anorexia, cancer-associated cachexia, cancer,  
 CC neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,  
 CC immune disorders, haematopoietic disorders, and the various  
 CC dyslipidaemias, metabolic disturbances associated with obesity, the  
 CC metabolic syndrome X and wasting disorders associated with chronic  
 CC diseases, bacterial, fungal, protozoal and viral infections, pain,

CC bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's  
CC disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina  
CC pectoris, myocardial infarction, ulcer, allergy, benign prostatic  
CC hypertrophy, and psychotic and neurological disorders, including anxiety,  
CC schizophrenia, manic depression, delirium, dementia, and dyskinesias,  
CC such as Huntington's disease and Gilles de la Tourette's syndrome. The  
CC nucleic acid is useful in gene therapy. The present sequence represents a  
CC NOVX protein

XX

SQ Sequence 898 AA;

Query Match 100.0%; Score 4660; DB 5; Length 898;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
          |||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy     61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL EEWYWCQVAVSSSGTTKSQKAYIRIA 120
          |||
Db     86 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGL EEWYWCQVAVSSSGTTKSQKAYIRIA 145

Qy    121 RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS 180
          |||
Db    146 RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS 205

Qy    181 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
          |||
Db    206 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

Qy    241 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECS DP 300
          |||
Db    266 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECS DP 325

Qy    301 APRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLI AVAVCLVLLLLVLILVYCR 360
          |||
Db    326 APRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLI AVAVCLVLLLLVLILVYCR 385

Qy    361 KKEGLDSDVADSSILTS GFQPVSIKPSKADNP HLLTIQPD LSTTTTTYQGS LCP RQDGPS 420
          |||
Db    386 KKEGLDSDVADSSILTS GFQPVSIKPSKADNP HLLTIQPD LSTTTTTYQGS LCP RQDGPS 445

Qy    421 PKFQLTN GHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQ NYFRSLPRGTSNMTYGT FNF 480
          |||
Db    446 PKFQLTN GHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQ NYFRSLPRGTSNMTYGT FNF 505

Qy    481 LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG 540
          |||
Db    506 LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG 565

Qy    541 VLLTRPVILAMDHCGEPS PDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACY 600
          |||
Db    566 VLLTRPVILAMDHCGEPS PDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACY 625

Qy    601 VFTEQLGRFALVGEALSVA AAKRLKLLL FAPVACTSLEYNIRVYCLHDTHDALKEVVQLE 660
          |||
```



Db 626 VFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE 685  
 Qy 661 KQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWWSKLLVSYQEIPFYHIWNGTQRYLH 720  
 |||||  
 Db 686 KQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWWSKLLVSYQEIPFYHIWNGTQRYLH 745  
 Qy 721 CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG 780  
 |||||  
 Db 746 CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG 805  
 Qy 781 PSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEA 840  
 |||||  
 Db 806 PSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEA 865  
 Qy 841 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 873  
 |||||  
 Db 866 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898

# RESULT 2

ADH71618

ID ADH71618 standard; protein; 898 AA.

XX

AC ADH71618;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21e SEQ ID NO:514.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.  
PR 10-JUN-2002; 2002US-0387400P.  
PR 10-JUN-2002; 2002US-0387535P.  
PR 11-JUN-2002; 2002US-0387610P.  
PR 11-JUN-2002; 2002US-0387625P.  
PR 11-JUN-2002; 2002US-0387634P.  
PR 11-JUN-2002; 2002US-0387668P.  
PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.



Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	180
Db	146	RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	205
Qy	181	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SR SCTNPAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECS DP	300
Db	266	SR SCTNPAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECS DP	325
Qy	301	APRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	360
Db	326	APRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	385
Qy	361	KKEGLDSDVADSSILTS GFQPVSIKPSKADNP HLLTIQPD LSTTTTTTYQGS LCP RQDGPS	420
Db	386	KKEGLDSDVADSSILTS GFQPVSIKPSKADNP HLLTIQPD LSTTTTTTYQGS LCP RQDGPS	445
Qy	421	PKFQLTN GHLLSPLGGGRHTLHHSSPTSEAE EFVSR LSTQNYFRSLPRGTSNM TYGT FNF	480
Db	446	PKFQLTN GHLLSPLGGGRHTLHHSSPTSEAE EFVSR LSTQNYFRSLPRGTSNM TYGT FNF	505
Qy	481	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG	540
Db	506	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG	565
Qy	541	VLLTRPVILAMDHCGEPS PDSWSLRLKKQSC EGSWEDVLHLGEEAPSHLYYCQLEASACY	600
Db	566	VLLTRPVILAMDHCGEPS PDSWSLRLKKQSC EGSWEDVLHLGEEAPSHLYYCQLEASACY	625
Qy	601	VFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	660
Db	626	VFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	685
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS LWSKLLVSYQEIPFYHIWNGTQRYLH	720
Db	686	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSS LWSKLLVSYQEIPFYHIWNGTQRYLH	745
Qy	721	CTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	780
Db	746	CTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	805
Qy	781	PSAFKIPFLIRQKIIS SLDPPCR RGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA	840
Db	806	PSAFKIPFLIRQKIIS SLDPPCR RGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA	865
Qy	841	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C	873
Db	866	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C	898

RESULT 3

AAU97899

ID AAU97899 standard; protein; 898 AA.

XX

AC AAU97899;  
 XX  
 DT 27-AUG-2002 (first entry)  
 XX  
 DE Human netrin binding membrane receptor UNC5H-1 protein.  
 XX  
 KW Netrin binding membrane receptor; receptor; UNC5H-1; human; nootropic;  
 KW neuroprotective; cytostatic; antiparkinsonian; cerebroprotective; cancer;  
 KW central nervous system; CNS; stroke; Parkinson's disease;  
 KW multiple sclerosis; Alzheimer's disease.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 152. .223  
 FT /note= "Immunoglobulin domain "  
 FT Domain 247. .294  
 FT /note= "Thrombospondine type 1 domain "  
 FT Domain 302. .348  
 FT /note= "Thrombospondine type 1 domain"  
 FT Region 361. .382  
 FT /note= "Transmembrane region"  
 FT Domain 495. .598  
 FT /note= "ZU5 domain"  
 FT Domain 817. .897  
 FT /note= "Death domain"  
 XX  
 PN WO200233080-A2.  
 XX  
 PD 25-APR-2002.  
 XX  
 PF 15-OCT-2001; 2001WO-EP011891.  
 XX  
 PR 16-OCT-2000; 2000US-0240061P.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Koehler RH;  
 XX  
 DR WPI; 2002-463314/49.  
 DR N-PSDB; ABK52891.  
 XX  
 PT Novel human netrin binding membrane receptor polypeptide and  
 PT polynucleotides for identifying modulating agents useful in treating  
 PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,  
 PT Alzheimer's disease.  
 XX  
 PS Claim 1; Fig 2; 94pp; English.  
 XX  
 CC This invention relates to the DNA and protein sequences of a novel  
 CC purified human netrin binding membrane receptor, UNC5H-1. The DNA  
 CC sequence of the invention is useful as a probe for detecting a nucleic  
 CC acid encoding the UNC5H-1 protein in a biological sample. The sequences  
 CC of the invention are useful to screen for agents which decrease the  
 CC activity of the UNC5H-1 protein. The sequences are also useful for  
 CC screening agents which regulate (modulate) the activity of the protein of  
 CC the invention. A pharmaceutical composition containing the protein of the

CC invention or a reagent that modulates the activity of the UNC5H-1 protein  
CC may be useful for treating a UNC5H-1 dysfunction related disease such as  
CC cancer or a central nervous system (CNS) disorders (e.g, Parkinson's  
CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion  
CC proteins comprising the UNC5H-1 protein are useful for generating  
CC antibodies and for in various assay systems, and the protein can be used  
CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method  
CC of the invention is useful for detecting a coding sequence for the UNC5H-  
CC 1 protein. The present sequence represents the human netrin binding  
CC membrane receptor UNC5H-1 protein of the invention

XX

SQ Sequence 898 AA;

Query Match 99.8%; Score 4650; DB 5; Length 898;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 871; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
      |||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy      61 DHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 120
      |||
Db      86 DHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 145

Qy     121 RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS 180
      |||
Db     146 YLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS 205

Qy     181 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
      |||
Db     206 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

Qy     241 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP 300
      |||
Db     266 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP 325

Qy     301 APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR 360
      |||
Db     326 APRNGGEECQGTDLDRNCTSDLCVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR 385

Qy     361 KKEGLDSDVADSSILTSQFQVSIKPSKADNPHELLTIQPDLSSTTTTQYQSLCPRQDGPS 420
      |||
Db     386 KKEGLDSDVADSSILTSQFQVSIKPSKADNPHELLTIQPDLSSTTTTQYQSLCPRQDGPS 445

Qy     421 PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFFVSRLSTQNYFRSLPRGTSNMTYGTFFNF 480
      |||
Db     446 PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFFVSRLSTQNYFRSLPRGTSNMTYGTFFNF 505

Qy     481 LGGRMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG 540
      |||
Db     506 LGGRMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG 565

Qy     541 VLLTRPVILAMDHCGEPSPDSSWLRLLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACY 600
      |||
Db     566 VLLTRPVILAMDHCGEPSPDSSWLRLLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACY 625
```

Qy	601	VFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	660
Db	626	VFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	685
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRYLH	720
Db	686	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRYLH	745
Qy	721	CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	780
Db	746	CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	805
Qy	781	PSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEA	840
Db	806	PSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEA	865
Qy	841	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	873
Db	866	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	898

RESULT 4

AAU79939

ID AAU79939 standard; protein; 899 AA.

XX

AC AAU79939;

XX

DT 15-JUL-2002 (first entry)

XX

DE Human UNC5-like protein NOV1.

XX

KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;  
 KW cell signal processing; metabolic pathway modulation; cancerous tissue;  
 KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;  
 KW chromosome 13.

XX

OS Homo sapiens.

XX

PN WO200229038-A2.

XX

PD 11-APR-2002.

XX

PF 04-OCT-2001; 2001WO-US031377.

XX

PR 04-OCT-2000; 2000US-0237862P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Herrmann JL, Rastelli L, Shimkets RA;

XX

DR WPI; 2002-340104/37.

DR

N-PSDB; ABK49422.

XX

PT Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for

PT

treating cardiomyopathy, artherosclerosis, and cancer.

XX

PS Claim 1; Page 9; 180pp; English.





Db	565	GVLLTRPVILAMDHCGEPSDWSLRLKKQSCGSEWDVHLHGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	658
Db	625	CYVFTEQLGRFALVGEALSVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	684
Qy	659	LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSCLLVSYQEIPFYHIWNGTQRY	718
Db	685	LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSCLLVSYQEIPFYHIWNGTQRY	744
Qy	719	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	778
Db	745	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	804
Qy	779	VGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLW	838
Db	805	VGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLW	864
Qy	839	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	873
Db	865	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899

RESULT 5

ADG42569

ID ADG42569 standard; protein; 899 AA.

XX

AC ADG42569;

XX

DT 26-FEB-2004 (first entry)

XX

DE Novel human NOV1.

XX

KW cytostatic; gene therapy; NOVX-agonist; NOVX-antagonist; pharmaceutical;

KW NOVX-associated disorder; cancer; human.

XX

OS Homo sapiens.

XX

PN US2003204052-A1.

XX

PD 30-OCT-2003.

XX

PF 04-OCT-2001; 2001US-00970944.

XX

PR 04-OCT-2000; 2000US-0237862P.

XX

PA (HERR/) HERRMANN J L.

PA (RAST/) RASTELLI L.

PA (SHIM/) SHIMKETS R A.

XX

PI Herrmann JL, Rastelli L, Shimkets RA;

XX

DR WPI; 2003-900673/82.

DR N-PSDB; ADG42568.

XX

PT New NOVX gene or NOVX-specific antibody, useful for preparing a

PT composition for treating or preventing a NOVX-associated disorder, e.g.,  
PT cancer.

XX

PS Claim 1; SEQ ID NO 2; 118pp; English.

XX

CC The invention describes a new isolated polypeptide comprising: a  
CC polypeptide or its mature form comprising a sequence not given in the  
CC specification; or a variant of (A), where one or more amino acid residues  
CC in the variant differs in no more than 15% from the amino acid sequence  
CC of the mature form. The pharmaceutical composition may be administered  
CC via oral, transdermal, rectal or parenteral route. The polypeptide,  
CC nucleic acid or antibody is useful for preparing a composition for  
CC treating or preventing a NOVX-associated disorder, e.g., cancer. This is  
CC the amino acid sequence of a human NOVX protein.

XX

SQ Sequence 899 AA;

Query Match 98.0%; Score 4567.5; DB 7; Length 899;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 863; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

Qy	1	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	60
Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSSGEPTMEVRINVSQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	180
Db	146	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	205
Qy	181	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SR SCTNPAPLNGGAFCEGQNV-QKTACATLCPVDGWSWSPWSKWSACGLDCTHWSRECS	299
		:  :	
Db	266	SR SCTNPAPLNGGAFCEGQNVHDTVSSLLVSDGWSWSPWSKWSACGLDCTHWSRECS	325
Qy	300	PAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLI AVAVCLV LLLVLILVYC	359
Db	326	PAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLI AVAVCLV LLLVLILVYC	385
Qy	360	RKKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPDLS-TT TTYQGSLCPRQDGP	419
Db	386	RKKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPDLS-TT TTYQGSLCPRQDGP	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNMTYGT FN	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNMTYGT FN	504
Qy	480	FLGGRLMIPNTGISLLIPPD AIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPPD AIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP	564

Qy 540 GVLLTRPVILAMDHCGEPSDWSLRLKKQSCGGSWE-DVLHLGEEAPSHLYYCQLEASA 598  
 |||||  
 Db 565 GVLLTRPVILAMDHCGEPSDWSLRLKKQSCGGSWEQDVLHLGEEAPSHLYYCQLEASA 624  
 |||||  
 Qy 599 CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ 658  
 |||||  
 Db 625 CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ 684  
 |||||  
 Qy 659 LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY 718  
 |||||  
 Db 685 LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY 744  
 |||||  
 Qy 719 LHCTFTLERVSPSTSDLACKLWWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL 778  
 |||||  
 Db 745 LHCTFTLERVSPSTSDLACKLWWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL 804  
 |||||  
 Qy 779 VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW 838  
 |||||  
 Db 805 VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW 864  
 |||||  
 Qy 839 EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C 873  
 |||||  
 Db 865 EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C 899  
 |||||

RESULT 6

ADH71610

ID ADH71610 standard; protein; 899 AA.

XX

AC ADH71610;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21a SEQ ID NO:506.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;  
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;  
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;  
 KW obesity; diabetes; infectious disease; metabolic syndrome X;  
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.  
PR 07-JUN-2002; 2002US-0386816P.  
PR 07-JUN-2002; 2002US-0386931P.  
PR 07-JUN-2002; 2002US-0386942P.  
PR 07-JUN-2002; 2002US-0386971P.  
PR 07-JUN-2002; 2002US-0387262P.  
PR 08-JUN-2002; 2002US-0296960P.  
PR 10-JUN-2002; 2002US-0387400P.  
PR 10-JUN-2002; 2002US-0387535P.  
PR 11-JUN-2002; 2002US-0387610P.  
PR 11-JUN-2002; 2002US-0387625P.  
PR 11-JUN-2002; 2002US-0387634P.  
PR 11-JUN-2002; 2002US-0387668P.  
PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.

PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.  
PR 31-OCT-2002; 2002US-0422690P.  
PR 01-NOV-2002; 2002US-0423130P.  
PR 05-NOV-2002; 2002US-00423798.  
PR 05-NOV-2002; 2002US-0423798P.  
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;  
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;  
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;  
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;  
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;  
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;  
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;  
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;  
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71609.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or  
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 506; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A  
CC polypeptide of the invention has cytostatic, immunomodulator,  
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and  
CC antilipaemic activity, and may have a use in gene therapy, and as a  
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising  
CC any of the 303 fully defined nucleotide sequences given in the  
CC specification. The polypeptide is useful in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease. The  
CC polypeptide, polynucleotide and antibody are useful in diagnosing,  
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,  
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious  
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are  
CC further used as hybridisation probes, in chromosome mapping, tissue ...  
CC typing, preventive medicine, and pharmacogenomics. The present sequence  
CC represents a NOVX polypeptide of the invention.

XX

SQ Sequence 899 AA;

Query Match 98.0%; Score 4567.5; DB 8; Length 899;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 863; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

Qy 1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60  
|||||

Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSSGEPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	180
Db	146	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	205
Qy	181	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SR SCTNPAPLNGGAFCEGQNV-QKTACATLCPVDGWSWPSKWSACGLDCTHWSRECS	299
		:  :	
Db	266	SR SCTNPAPLNGGAFCEGQNVHDRTVSSLLVSV DGSWSPWSKWSACGLDCTHWSRECS	325
Qy	300	PAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLV LLLLVLILVYC	359
Db	326	PAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLV LLLLVLILVYC	385
Qy	360	RKKEGLSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LSTTTTTYQGS LCP RQDGP	419
Db	386	RKKEGLSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LS-TTTYQGS LCP RQDGP	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FN	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FN	504
Qy	480	FLGGRLMIPNTGISLLIPDAIPRGKIYEIY LTLHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPDAIPRGKIYEIY LTLHKPEDVRLPLAGCQTLLSPIVSCGPP	564
Qy	540	GVLLTRPVILAMDHCGEPS PDSWSLRLKKQSC EG SWE-DVLHLGEEAPSHLYYCQLEASA	598
Db	565	GVLLTRPVILAMDHCGEPS PDSWSLRLKKQSC EG SWEQDVLHLGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	658
Db	625	CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	684
Qy	659	LEKQLGGQLIQEP RVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY	718
Db	685	LEKQLGGQLIQEP RVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY	744
Qy	719	LHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	778
Db	745	LHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	804
Qy	779	VGPSAFKIPFLIRQKIIS SLDPPCR RGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW	838
Db	805	VGPSAFKIPFLIRQKIIS SLDPPCR RGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW	864
Qy	839	EARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAE C	873
Db	865	EARHFPNGNLSQLAAA VAGLGQPDAGLFTVSEAE C	899

RESULT 7

ADH71628

ID ADH71628 standard; protein; 899 AA.

XX

AC ADH71628;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21j SEQ ID NO:524.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0389123P.

PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.  
PR 31-OCT-2002; 2002US-0422690P.  
PR 01-NOV-2002; 2002US-0423130P.  
PR 05-NOV-2002; 2002US-0423798.  
PR 05-NOV-2002; 2002US-0423798P.  
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;  
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;  
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;  
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;  
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;  
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;  
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;  
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;  
PI Zhong H;



XX WPI; 2004-081935/08.  
DR N-PSDB; ADH71627.  
XX  
PT New NOVX polypeptides and nucleic acid molecules useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or  
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.  
XX  
PS Example 21; SEQ ID NO 524; 1880pp; English.  
XX  
CC The invention relates to a novel isolated polypeptide (NOVX). A  
CC polypeptide of the invention has cytostatic, immunomodulator,  
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and  
CC antilipaemic activity, and may have a use in gene therapy, and as a  
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising  
CC any of the 303 fully defined nucleotide sequences given in the  
CC specification. The polypeptide is useful in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease. The  
CC polypeptide, polynucleotide and antibody are useful in diagnosing,  
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,  
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious  
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are  
CC further used as hybridisation probes, in chromosome mapping, tissue  
CC typing, preventive medicine, and pharmacogenomics. The present sequence  
CC represents a NOVX polypeptide of the invention.  
XX  
SQ Sequence 899 AA;



XX  
PF 03-JUN-2003; 2003WO-US017430.  
XX  
PR 03-JUN-2002; 2002US-0385120P.  
PR 04-JUN-2002; 2002US-0385784P.  
PR 05-JUN-2002; 2002US-0386041P.  
PR 05-JUN-2002; 2002US-0386047P.  
PR 06-JUN-2002; 2002US-0386376P.  
PR 06-JUN-2002; 2002US-0386453P.  
PR 06-JUN-2002; 2002US-0386864P.  
PR 06-JUN-2002; 2002US-0387016P.  
PR 07-JUN-2002; 2002US-0386796P.  
PR 07-JUN-2002; 2002US-0386816P.  
PR 07-JUN-2002; 2002US-0386931P.  
PR 07-JUN-2002; 2002US-0386942P.  
PR 07-JUN-2002; 2002US-0386971P.  
PR 07-JUN-2002; 2002US-0387262P.  
PR 08-JUN-2002; 2002US-0296960P.  
PR 10-JUN-2002; 2002US-0387400P.  
PR 10-JUN-2002; 2002US-0387535P.  
PR 11-JUN-2002; 2002US-0387610P.  
PR 11-JUN-2002; 2002US-0387625P.  
PR 11-JUN-2002; 2002US-0387634P.  
PR 11-JUN-2002; 2002US-0387668P.  
PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.

PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.  
PR 31-OCT-2002; 2002US-0422690P.  
PR 01-NOV-2002; 2002US-0423130P.  
PR 05-NOV-2002; 2002US-00423798.  
PR 05-NOV-2002; 2002US-0423798P.  
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;  
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;  
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;  
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;  
PI MacLachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;  
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;  
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;  
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;  
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71635.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or  
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 532; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A  
CC polypeptide of the invention has cytostatic, immunomodulator,  
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and  
CC antilipaemic activity, and may have a use in gene therapy, and as a  
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising  
CC any of the 303 fully defined nucleotide sequences given in the  
CC specification. The polypeptide is useful in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease. The  
CC polypeptide, polynucleotide and antibody are useful in diagnosing,  
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,  
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious  
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are  
CC further used as hybridisation probes, in chromosome mapping, tissue

CC typing, preventive medicine, and pharmacogenomics. The present sequence  
CC represents a NOVX polypeptide of the invention.  
XX  
SQ Sequence 899 AA;

Query Match 98.0%; Score 4566.5; DB 8; Length 899;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 862; Conservative 3; Mismatches 7; Indels 3; Gaps 3;

Qy	1	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	60
Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVSROQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSSGEPTMEVRINVSROQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDPSLDPNVYITREHS	180
Db	146	RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDPSLDPNVYITREHS	205
Qy	181	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS	299
		:  :	
Db	266	SRSCTNPAPLNGGAFCEGQNVHDRVSSLLVSVDGSWSPWSKWSACGLDCTHWSRECS	325
Qy	300	PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIYVAVCLVLLLLVLILVYC	359
Db	326	PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIYVAVCLVLLLLVLILVYC	385
Qy	360	RKKEGLDSDVADSSILTSQFQVSIKPSKADNPHLLTIQPDLS-TTTTYQGSCLPRQDGP	419
Db	386	RKKEGLDSDVADSSILTSQFQVSIKPSKADNPHLLTIQPDLS-TTTTYQGSCLPRQDGP	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMITYGT	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMITYGT	504
Qy	480	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPP	564
Qy	540	GVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWE-DVLHLGEEAPSHLYYCQLEASA	598
Db	565	GVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVAALKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEV	658
		:	
Db	625	CYIFTEQLGRFALVGEALSVAALKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEV	684
Qy	659	LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDPVSSLWKSLLVSYQEIPFYHIWNGTQ	718
Db	685	LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDPVSSLWKSLLVSYQEIPFYHIWNGTQ	744

Qy 719 LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL 778  
 Db 745 LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL 804  
 Qy 779 VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW 838  
 Db 805 VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW 864  
 Qy 839 EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 873  
 Db 865 EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899

RESULT 9

ADH71642

ID ADH71642 standard; protein; 899 AA.

XX

AC ADH71642;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21q SEQ ID NO:538.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;  
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;  
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;  
 KW obesity; diabetes; infectious disease; metabolic syndrome X;  
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.  
PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.  
PR 31-OCT-2002; 2002US-0422690P.  
PR 01-NOV-2002; 2002US-0423130P.  
PR 05-NOV-2002; 2002US-00423798.  
PR 05-NOV-2002; 2002US-0423798P.  
PR 12-NOV-2002; 2002US-0425453P.  
XX





Db	206	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SR SCTNPAPLNGGAFCEGQNV-QKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECS	299
		:  :	
Db	266	SR SCTNPAPLNGGAFCEGQNVHDRVSSLLVSDGWSWPWSKWSACGLDCTHWSRECS	325
Qy	300	PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLI AVAVCLVLLLLVLILVYC	359
Db	326	PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLI AVAVCLVLLLLVLILVYC	385
Qy	360	RKKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPDLS TTTTYQGS LCPRQDGP	419
Db	386	RKKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPDLS -TTTTYQGS LCPRQDGP	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FN	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FN	504
Qy	480	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP	564
Qy	540	GVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWE-DVLHLGEEAPSHLYYCQLEASA	598
Db	565	GVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	658
		:	
Db	625	CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKELVQ	684
Qy	659	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY	718
Db	685	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY	744
Qy	719	LHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	778
Db	745	LHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	804
Qy	779	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW	838
Db	805	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW	864
Qy	839	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	873
Db	865	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	899

RESULT 10

ADH71640

ID ADH71640 standard; protein; 899 AA.

XX

AC ADH71640;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21p SEQ ID NO:536.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;  
KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;  
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;  
KW obesity; diabetes; infectious disease; metabolic syndrome X;  
KW dyslipidaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO2003102155-A2.  
XX  
PD 11-DEC-2003.  
XX  
PF 03-JUN-2003; 2003WO-US017430.  
XX  
PR 03-JUN-2002; 2002US-0385120P.  
PR 04-JUN-2002; 2002US-0385784P.  
PR 05-JUN-2002; 2002US-0386041P.  
PR 05-JUN-2002; 2002US-0386047P.  
PR 06-JUN-2002; 2002US-0386376P.  
PR 06-JUN-2002; 2002US-0386453P.  
PR 06-JUN-2002; 2002US-0386864P.  
PR 06-JUN-2002; 2002US-0387016P.  
PR 07-JUN-2002; 2002US-0386796P.  
PR 07-JUN-2002; 2002US-0386816P.  
PR 07-JUN-2002; 2002US-0386931P.  
PR 07-JUN-2002; 2002US-0386942P.  
PR 07-JUN-2002; 2002US-0386971P.  
PR 07-JUN-2002; 2002US-0387262P.  
PR 08-JUN-2002; 2002US-0296960P.  
PR 10-JUN-2002; 2002US-0387400P.  
PR 10-JUN-2002; 2002US-0387535P.  
PR 11-JUN-2002; 2002US-0387610P.  
PR 11-JUN-2002; 2002US-0387625P.  
PR 11-JUN-2002; 2002US-0387634P.  
PR 11-JUN-2002; 2002US-0387668P.  
PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.

PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.  
PR 31-OCT-2002; 2002US-0422690P.  
PR 01-NOV-2002; 2002US-0423130P.  
PR 05-NOV-2002; 2002US-00423798.  
PR 05-NOV-2002; 2002US-0423798P.  
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;  
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;  
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;  
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;  
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;  
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;  
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;  
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;  
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71639.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or  
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 536; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A  
CC polypeptide of the invention has cytostatic, immunomodulator,

CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and  
 CC antilipaemic activity, and may have a use in gene therapy, and as a  
 CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising  
 CC any of the 303 fully defined nucleotide sequences given in the  
 CC specification. The polypeptide is useful in the manufacture of a  
 CC medicament for treating a syndrome associated with a human disease. The  
 CC polypeptide, polynucleotide and antibody are useful in diagnosing,  
 CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,  
 CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious  
 CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are  
 CC further used as hybridisation probes, in chromosome mapping, tissue  
 CC typing, preventive medicine, and pharmacogenomics. The present sequence  
 CC represents a NOVX polypeptide of the invention.

XX

SQ Sequence 899 AA;

Query Match 98.0%; Score 4564.5; DB 8; Length 899;  
 Best Local Similarity 98.5%; Pred. No. 0;  
 Matches 862; Conservative 3; Mismatches 7; Indels 3; Gaps 3;

Qy	1	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	60
Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVSROQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSSGEPTMEVRINVSROQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	180
Db	146	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	205
Qy	181	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SR SCTNPAPLNGGAFCEGQNV-QKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS	299
		:  :	
Db	266	SR SCTNPAPLNGGAFCEGQNVHDTVSSLLVSVDGSWSPWSKWSACGLDCTHWSRECS	325
Qy	300	PAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLI AVAVCLV LLLVLILVYC	359
Db	326	PAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLI AVAVCLV LLLVLILVYC	385
Qy	360	RKKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LSTTTT TYQGS LCP RQDGP	419
Db	386	RKKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LSTTTT TYQGS LCP RQDGP	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FN	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FN	504
Qy	480	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP	564
Qy	540	GVLLTRPVILAMDHCGEPSDWSLRLKKQSCGSWE-DVLHLGEEAPSHLYYCQLEASA	598

Db	565	GVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	658
Db	625	CYVFTEQLGRFALVGEALSVA AAKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	684
Qy	659	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY	718
Db	685	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY	744
Qy	719	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	778
Db	745	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	804
Qy	779	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW	838
Db	805	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW	864
Qy	839	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	873
Db	865	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899

RESULT 11

ADH71632

ID ADH71632 standard; protein; 899 AA.

XX

AC ADH71632;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV211 SEQ ID NO:528.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;  
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;  
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;  
 KW obesity; diabetes; infectious disease; metabolic syndrome X;  
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.  
PR 07-JUN-2002; 2002US-0386931P.  
PR 07-JUN-2002; 2002US-0386942P.  
PR 07-JUN-2002; 2002US-0386971P.  
PR 07-JUN-2002; 2002US-0387262P.  
PR 08-JUN-2002; 2002US-0296960P.  
PR 10-JUN-2002; 2002US-0387400P.  
PR 10-JUN-2002; 2002US-0387535P.  
PR 11-JUN-2002; 2002US-0387610P.  
PR 11-JUN-2002; 2002US-0387625P.  
PR 11-JUN-2002; 2002US-0387634P.  
PR 11-JUN-2002; 2002US-0387668P.  
PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.

XX

XX

XX

DR

XX

XX

XX

XX

Qy 1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLLVCKAVPATQIFFKCNGEWEVRQV 60  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLLVCKAVPATQIFFKCNGEWEVROV 85

Qy	61	DHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSSGSEPTMEVRINVSRRQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS	180
Db	146	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS	205
Qy	181	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVQRARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDGWSWPSKWSACGLDCTHWSRECS	299
		:  :	
Db	266	SRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSDGWSWPSKWSACGLDCTHWSRECS	325
Qy	300	PAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYC	359
Db	326	PAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYC	385
Qy	360	RKKEGLSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQGSLCPRQDGP	419
Db	386	RKKEGLSDVADSSILTSGFQPVSIKPSKADNPHLLTIQPDLS-TTTTYQGSLCPRQDGP	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFN	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFN	504
Qy	480	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPDVIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPP	564
Qy	540	GVLLTRPVILAMDHCGEPSPDWSLRLKKQSCGSWE-DVLHLGEEAPSHLYYCQLEASA	598
Db	565	GVLLTRPVILAMDHCGEPSPDWSLRLKKQSCGSWEQDVLHLGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVAAPAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	658
Db	625	CYVFTEQLGRFALVGEALSVAAPAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	684
Qy	659	LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRY	718
Db	685	LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRY	744
Qy	719	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	778
Db	745	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	804
Qy	779	VGPSAFKIPFLIRQKIISLDPPCRGGADWRTLAQKLHLDHLSFFASKPSPTAMILNLW	838
Db	805	VGPSAFKIPFLIRQKIISLDPPCRGGADWRTLAQKLHLDHLSFFASKPSPTAMILNLW	864
Qy	839	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	873
Db	865	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899



RESULT 12

ADH71638

ID ADH71638 standard; protein; 899 AA.

XX

AC ADH71638;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21o SEQ ID NO:534.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.

PR 11-JUN-2002; 2002US-0387696P.

PR 11-JUN-2002; 2002US-0387702P.

PR 11-JUN-2002; 2002US-0387836P.

PR 11-JUN-2002; 2002US-0387859P.

PR 12-JUN-2002; 2002US-0387933P.

PR 12-JUN-2002; 2002US-0387934P.

PR 12-JUN-2002; 2002US-0387960P.

PR 12-JUN-2002; 2002US-0388022P.

PR 12-JUN-2002; 2002US-0388096P.

PR 13-JUN-2002; 2002US-0389123P.

PR 14-JUN-2002; 2002US-0389118P.

PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.  
PR 31-OCT-2002; 2002US-0422690P.  
PR 01-NOV-2002; 2002US-0423130P.  
PR 05-NOV-2002; 2002US-00423798.  
PR 05-NOV-2002; 2002US-0423798P.  
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;  
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;  
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;  
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;  
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;  
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;  
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;  
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;  
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71637.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or  
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 534; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A  
CC polypeptide of the invention has cytostatic, immunomodulator,  
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and  
CC antilipaemic activity, and may have a use in gene therapy, and as a  
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising  
CC any of the 303 fully defined nucleotide sequences given in the  
CC specification. The polypeptide is useful in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease. The  
CC polypeptide, polynucleotide and antibody are useful in diagnosing,  
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,  
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious  
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are  
CC further used as hybridisation probes, in chromosome mapping, tissue  
CC typing, preventive medicine, and pharmacogenomics. The present sequence  
CC represents a NOVX polypeptide of the invention.

XX

SQ Sequence 899 AA;

Query Match 97.9%; Score 4563.5; DB 8; Length 899;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 862; Conservative 2; Mismatches 8; Indels 3; Gaps 3;

```
Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
          |||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy     61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEWYWCQCVAWSSSGTTKSQKAYIRIA 120
          |||
Db     86 DHVIERSTDGSSGEPTMEVRINVSRRQVEKVFGLLEEWYWCQCVAWSSSGTTKSQKAYIRIA 145

Qy    121 RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS 180
          |||
Db    146 RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS 205

Qy    181 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
          |||
Db    206 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

Qy    241 SRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS 299
          ||| :| :| |||
Db    266 SRSCTNPAPLNGGAFCEGQNVHDRVSSLLVSVDGSWSPWSKWSACGLDCTHWSRECS 325

Qy    300 PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYC 359
          |||
Db    326 PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYC 385

Qy    360 RKKEGLSDVDADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSSTTTTTYQGSCLPRQDGP 419
          |||
```

Db	386	RKKEGLDSDVADSSILTSQFQVSIKPSKADNPHLLTIQPDLS-TTTTYQGSILCPRQDGP	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFFN	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFFN	504
Qy	480	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLTLLHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLTLLHKPEDVRLPLAGCQTLLSPIVSCGPP	564
Qy	540	GVLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWE-DVLHLGEEAPSHLYYCQLEASA	598
Db	565	GVLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVAALKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	658
Db	625	CYVFTEQLGRFALVGEALSVAATKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	684
Qy	659	LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRY	718
Db	685	LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRY	744
Qy	719	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	778
Db	745	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	804
Qy	779	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLW	838
Db	805	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLW	864
Qy	839	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	873
Db	865	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	899

# RESULT 13

ADH71634

ID ADH71634 standard; protein; 899 AA.

XX

AC ADH71634;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21m SEQ ID NO:530.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;  
 KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;  
 KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;  
 KW obesity; diabetes; infectious disease; metabolic syndrome X;  
 KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.  
XX  
PR 03-JUN-2002; 2002US-0385120P.  
PR 04-JUN-2002; 2002US-0385784P.  
PR 05-JUN-2002; 2002US-0386041P.  
PR 05-JUN-2002; 2002US-0386047P.  
PR 06-JUN-2002; 2002US-0386376P.  
PR 06-JUN-2002; 2002US-0386453P.  
PR 06-JUN-2002; 2002US-0386864P.  
PR 06-JUN-2002; 2002US-0387016P.  
PR 07-JUN-2002; 2002US-0386796P.  
PR 07-JUN-2002; 2002US-0386816P.  
PR 07-JUN-2002; 2002US-0386931P.  
PR 07-JUN-2002; 2002US-0386942P.  
PR 07-JUN-2002; 2002US-0386971P.  
PR 07-JUN-2002; 2002US-0387262P.  
PR 08-JUN-2002; 2002US-0296960P.  
PR 10-JUN-2002; 2002US-0387400P.  
PR 10-JUN-2002; 2002US-0387535P.  
PR 11-JUN-2002; 2002US-0387610P.  
PR 11-JUN-2002; 2002US-0387625P.  
PR 11-JUN-2002; 2002US-0387634P.  
PR 11-JUN-2002; 2002US-0387668P.  
PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.

PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.  
PR 31-OCT-2002; 2002US-0422690P.  
PR 01-NOV-2002; 2002US-0423130P.  
PR 05-NOV-2002; 2002US-00423798.  
PR 05-NOV-2002; 2002US-0423798P.  
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;  
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;  
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;  
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;  
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;  
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;  
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;  
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;  
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71633.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or  
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 530; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A  
CC polypeptide of the invention has cytostatic, immunomodulator,  
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and  
CC antilipaemic activity, and may have a use in gene therapy, and as a  
CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising  
CC any of the 303 fully defined nucleotide sequences given in the  
CC specification. The polypeptide is useful in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease. The  
CC polypeptide, polynucleotide and antibody are useful in diagnosing,  
CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,  
CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious  
CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are  
CC further used as hybridisation probes, in chromosome mapping, tissue  
CC typing, preventive medicine, and pharmacogenomics. The present sequence

CC represents a NOVX polypeptide of the invention.

XX

SQ Sequence 899 AA;

Query Match 97.9%; Score 4562.5; DB 8; Length 899;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 862; Conservative 2; Mismatches 8; Indels 3; Gaps 3;

```
Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
          |||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy     61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 120
          |||
Db     86 DHVIERSTDGSSGPEPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 145

Qy    121 RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAAEVEWLRNEDLVDPSLDPNVYITREHS 180
          |||
Db    146 RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAAEVEWLRNEDLVDPSLDPNVYITREHS 205

Qy    181 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
          |||
Db    206 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

Qy    241 SRSCNPAPLNGGAFCEGQNV-QKTACATLCFVDGWSWPSKWSACGLDCTHWSRECS 299
          |||
Db    266 SRSCNPAPLNGGAFCEGQNVHDRTVSSLLVSDGWSWPSKWSACGLDCTHWSRECS 325

Qy    300 PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIYVAVCLVLLLLVLILVYC 359
          |||
Db    326 PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIYVAVCLVLLLLVLILVYC 385

Qy    360 RKKEGLDSDVADSSILTSFGFQVSIKPSKADNPHLLTIQPDLS-TTTTYQGSLCPRQDGP 419
          |||
Db    386 RKKEGLDSDVADSSILTSFGFQVSIKPSKADNPHLLTIQPDLS-TTTTYQGSLCPRQDGP 444

Qy    420 SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMITYGTFN 479
          |||
Db    445 SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMITYGTFN 504

Qy    480 FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP 539
          |||
Db    505 FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP 564

Qy    540 GVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWE-DVLHLGEEAPSHLYYCQLEASA 598
          |||
Db    565 GVLLTRPVILAMDHCGEPSPDWSLHLKKQSCEGSWEQDVLHLGEEAPSHLYYCQLEASA 624

Qy    599 CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ 658
          |||
Db    625 CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ 684

Qy    659 LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDPSSLWKS KLLVSYQEIPFYHIWNGTQRY 718
          |||
Db    685 LEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDPSSLWKS KLLVSYQEIPFYHIWNGTQRY 744

Qy    719 LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL 778
```

```

Db      745 LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL 804
Qy      779 VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLW 838
Db      805 VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLW 864
Qy      839 EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 873
Db      865 EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 899

```

RESULT 14

ADH71648

ID ADH71648 standard; protein; 899 AA.

XX

AC ADH71648;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21t SEQ ID NO:544.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;

KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;

KW obesity; diabetes; infectious disease; metabolic syndrome X;

KW dyslipidaemia.

XX

OS Homo sapiens.

XX

PN WO2003102155-A2.

XX

PD 11-DEC-2003.

XX

PF 03-JUN-2003; 2003WO-US017430.

XX

PR 03-JUN-2002; 2002US-0385120P.

PR 04-JUN-2002; 2002US-0385784P.

PR 05-JUN-2002; 2002US-0386041P.

PR 05-JUN-2002; 2002US-0386047P.

PR 06-JUN-2002; 2002US-0386376P.

PR 06-JUN-2002; 2002US-0386453P.

PR 06-JUN-2002; 2002US-0386864P.

PR 06-JUN-2002; 2002US-0387016P.

PR 07-JUN-2002; 2002US-0386796P.

PR 07-JUN-2002; 2002US-0386816P.

PR 07-JUN-2002; 2002US-0386931P.

PR 07-JUN-2002; 2002US-0386942P.

PR 07-JUN-2002; 2002US-0386971P.

PR 07-JUN-2002; 2002US-0387262P.

PR 08-JUN-2002; 2002US-0296960P.

PR 10-JUN-2002; 2002US-0387400P.

PR 10-JUN-2002; 2002US-0387535P.

PR 11-JUN-2002; 2002US-0387610P.

PR 11-JUN-2002; 2002US-0387625P.

PR 11-JUN-2002; 2002US-0387634P.

PR 11-JUN-2002; 2002US-0387668P.



PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.  
PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.  
PR 31-OCT-2002; 2002US-0422690P.  
PR 01-NOV-2002; 2002US-0423130P.  
PR 05-NOV-2002; 2002US-00423798.  
PR 05-NOV-2002; 2002US-0423798P.  
PR 12-NOV-2002; 2002US-0425453P.  
XX  
PA (CURA-) CURAGEN CORP.

```

XX
PI  Alsobrook JP,  Alvarez E,  Anderson DW,  Boldog FL,  Casman SJ;
PI  Catterton E,  Chapoval A,  Crabtree-Bokor JR,  Edinger SR,  Ellerman K;
PI  Ettenberg S,  Gangolli EA,  Gerlach VL,  Gorman L,  Gunther E,  Guo X;
PI  Gusev VY,  Herrmann JL,  Ji W,  Kekuda R,  Li L,  Liu X,  Macdougall JR;
PI  Maclachlan T,  Malyankar UM,  Mezick AJ,  Millet I,  Mishra VS;
PI  Padigar M,  Patturajan M,  Pena CEA,  Peyman JA,  Raha D,  Rastelli L;
PI  Rieger DK,  Rothenberg ME,  Sciore P,  Shenoy SG,  Shimkets RA;
PI  Smithson G,  Spytek KA,  Stone DJ,  Vernet CAM,  Voss EZ,  Zhong M;
PI  Zhong H;
XX
DR  WPI; 2004-081935/08.
DR  N-PSDB; ADH71647.
XX
PT  New NOVX polypeptides and nucleic acid molecules useful for preventing or
PT  treating NOVX-associated disorders, e.g. cancer, diabetes, infection or
PT  obesity, and in chromosome mapping, tissue typing or pharmacogenomics.
XX
PS  Example 21; SEQ ID NO 544; 1880pp; English.
XX
CC  The invention relates to a novel isolated polypeptide (NOVX). A
CC  polypeptide of the invention has cytostatic, immunomodulator,
CC  neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and
CC  antilipaemic activity, and may have a use in gene therapy, and as a
CC  vaccine. The polypeptides are encoded by NOVX polynucleotides comprising
CC  any of the 303 fully defined nucleotide sequences given in the
CC  specification. The polypeptide is useful in the manufacture of a
CC  medicament for treating a syndrome associated with a human disease. The
CC  polypeptide, polynucleotide and antibody are useful in diagnosing,
CC  treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,
CC  Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious
CC  diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are
CC  further used as hybridisation probes, in chromosome mapping, tissue
CC  typing, preventive medicine, and pharmacogenomics. The present sequence
CC  represents a NOVX polypeptide of the invention.
XX
SQ  Sequence 899 AA;

Query Match          97.9%; Score 4560.5; DB 8; Length 899;
Best Local Similarity 98.5%; Pred. No. 0;
Matches 862; Conservative 2; Mismatches 8; Indels 3; Gaps 3;

Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy      61 DHVIERSTDGSSGLPTMEVRINVSRRQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 120
Db      86 DHVIERSTDGSSGPEPTMEVRINVSRRQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 145

Qy     121 RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDSLDPNVYITREHS 180
Db     146 RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDSLDPNVYITREHS 205

Qy     181 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
Db     206 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

```

Qy	241	SR SCTNPAPLNGGAFCEGQNV-QKTACATLC PVDGSPWSKWSACGLDCTHWSRECS	299
		:	
Db	266	SR SCTNPAPLNGGAFCEGQNVH DRTVSSLLVSVDGSPWSKWSACGLDCTHWSRECS	325
Qy	300	PAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYC	359
Db	326	PAPRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYC	385
Qy	360	RKKEGLDS D VADSSILTS GFQPVSIKPSKADNPHLLTIQPDLS TTTT TYQGS LCPRQDGP	419
Db	386	RKKEGLDS D VADSSILTS GFQPVSIKPSKADNPHLLTIQPDLS -TTT TYQGS LCPRQDGP	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FN	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FN	504
Qy	480	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPP	564
Qy	540	GVLLTRPVILAMDHC GEPSDWSLRLKKQSC EGSW E-DVLHLGEEAPSHLYYCQLEASA	598
Db	565	GVLLTRPVILAMDHC GEPSDWSLRLKKQSC EGSW EQDVLHLGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	658
Db	625	CYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	684
Qy	659	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY	718
Db	685	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRY	744
Qy	719	LHCTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	778
Db	745	LHCTFTLERVSPSTSD LACKLWVWQVEGGGQSFSINFNITKDTRFAELLALESEAGVPAL	804
Qy	779	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW	838
Db	805	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLW	864
Qy	839	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C	873
Db	865	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C	899

RESULT 15

ADH71630

ID ADH71630 standard; protein; 899 AA.

XX

AC ADH71630;

XX

DT 25-MAR-2004 (first entry)

XX

DE Human protein of the invention NOV21k SEQ ID NO:526.

XX

KW human; cytostatic; immunomodulator; neuroprotective; nootropic;

KW anorectic; antidiabetic; antimicrobial; antilipaemic; gene therapy;  
KW vaccine; cancer; cachexia; Alzheimer's disease; Parkinson's disease;  
KW obesity; diabetes; infectious disease; metabolic syndrome X;  
KW dyslipidaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO2003102155-A2.  
XX  
PD 11-DEC-2003.  
XX  
PF 03-JUN-2003; 2003WO-US017430.  
XX  
PR 03-JUN-2002; 2002US-0385120P.  
PR 04-JUN-2002; 2002US-0385784P.  
PR 05-JUN-2002; 2002US-0386041P.  
PR 05-JUN-2002; 2002US-0386047P.  
PR 06-JUN-2002; 2002US-0386376P.  
PR 06-JUN-2002; 2002US-0386453P.  
PR 06-JUN-2002; 2002US-0386864P.  
PR 06-JUN-2002; 2002US-0387016P.  
PR 07-JUN-2002; 2002US-0386796P.  
PR 07-JUN-2002; 2002US-0386816P.  
PR 07-JUN-2002; 2002US-0386931P.  
PR 07-JUN-2002; 2002US-0386942P.  
PR 07-JUN-2002; 2002US-0386971P.  
PR 07-JUN-2002; 2002US-0387262P.  
PR 08-JUN-2002; 2002US-0296960P.  
PR 10-JUN-2002; 2002US-0387400P.  
PR 10-JUN-2002; 2002US-0387535P.  
PR 11-JUN-2002; 2002US-0387610P.  
PR 11-JUN-2002; 2002US-0387625P.  
PR 11-JUN-2002; 2002US-0387634P.  
PR 11-JUN-2002; 2002US-0387668P.  
PR 11-JUN-2002; 2002US-0387696P.  
PR 11-JUN-2002; 2002US-0387702P.  
PR 11-JUN-2002; 2002US-0387836P.  
PR 11-JUN-2002; 2002US-0387859P.  
PR 12-JUN-2002; 2002US-0387933P.  
PR 12-JUN-2002; 2002US-0387934P.  
PR 12-JUN-2002; 2002US-0387960P.  
PR 12-JUN-2002; 2002US-0388022P.  
PR 12-JUN-2002; 2002US-0388096P.  
PR 13-JUN-2002; 2002US-0389123P.  
PR 14-JUN-2002; 2002US-0389118P.  
PR 14-JUN-2002; 2002US-0389120P.  
PR 14-JUN-2002; 2002US-0389144P.  
PR 14-JUN-2002; 2002US-0389146P.  
PR 17-JUN-2002; 2002US-0389729P.  
PR 17-JUN-2002; 2002US-0389742P.  
PR 18-JUN-2002; 2002US-0389884P.  
PR 19-JUN-2002; 2002US-0390006P.  
PR 19-JUN-2002; 2002US-0390209P.  
PR 21-JUN-2002; 2002US-0390763P.  
PR 17-JUL-2002; 2002US-0396706P.  
PR 06-AUG-2002; 2002US-0401628P.  
PR 09-AUG-2002; 2002US-0402156P.

PR 09-AUG-2002; 2002US-0402256P.  
PR 09-AUG-2002; 2002US-0402389P.  
PR 12-AUG-2002; 2002US-0402786P.  
PR 12-AUG-2002; 2002US-0402816P.  
PR 12-AUG-2002; 2002US-0402821P.  
PR 12-AUG-2002; 2002US-0402832P.  
PR 13-AUG-2002; 2002US-0403448P.  
PR 13-AUG-2002; 2002US-0403459P.  
PR 13-AUG-2002; 2002US-0403531P.  
PR 13-AUG-2002; 2002US-0403532P.  
PR 13-AUG-2002; 2002US-0403563P.  
PR 13-AUG-2002; 2002US-0406317P.  
PR 15-AUG-2002; 2002US-0403617P.  
PR 26-AUG-2002; 2002US-0406182P.  
PR 26-AUG-2002; 2002US-0406355P.  
PR 27-AUG-2002; 2002US-0406240P.  
PR 12-SEP-2002; 2002US-0410084P.  
PR 20-SEP-2002; 2002US-0412528P.  
PR 23-SEP-2002; 2002US-0412731P.  
PR 30-SEP-2002; 2002US-0414801P.  
PR 30-SEP-2002; 2002US-0414839P.  
PR 30-SEP-2002; 2002US-0414840P.  
PR 30-SEP-2002; 2002US-0414954P.  
PR 09-OCT-2002; 2002US-0417186P.  
PR 09-OCT-2002; 2002US-0417406P.  
PR 23-OCT-2002; 2002US-0420639P.  
PR 28-OCT-2002; 2002US-0421156P.  
PR 31-OCT-2002; 2002US-0422690P.  
PR 01-NOV-2002; 2002US-0423130P.  
PR 05-NOV-2002; 2002US-0423798.  
PR 05-NOV-2002; 2002US-0423798P.  
PR 12-NOV-2002; 2002US-0425453P.

XX

PA (CURA-) CURAGEN CORP.

XX

PI Alsobrook JP, Alvarez E, Anderson DW, Boldog FL, Casman SJ;  
PI Catterton E, Chapoval A, Crabtree-Bokor JR, Edinger SR, Ellerman K;  
PI Ettenberg S, Gangolli EA, Gerlach VL, Gorman L, Gunther E, Guo X;  
PI Gusev VY, Herrmann JL, Ji W, Kekuda R, Li L, Liu X, Macdougall JR;  
PI Maclachlan T, Malyankar UM, Mezick AJ, Millet I, Mishra VS;  
PI Padigar M, Patturajan M, Pena CEA, Peyman JA, Raha D, Rastelli L;  
PI Rieger DK, Rothenberg ME, Sciore P, Shenoy SG, Shimkets RA;  
PI Smithson G, Spytek KA, Stone DJ, Vernet CAM, Voss EZ, Zhong M;  
PI Zhong H;

XX

DR WPI; 2004-081935/08.

DR N-PSDB; ADH71629.

XX

PT New NOVX polypeptides and nucleic acid molecules useful for preventing or  
PT treating NOVX-associated disorders, e.g. cancer, diabetes, infection or  
PT obesity, and in chromosome mapping, tissue typing or pharmacogenomics.

XX

PS Example 21; SEQ ID NO 526; 1880pp; English.

XX

CC The invention relates to a novel isolated polypeptide (NOVX). A  
CC polypeptide of the invention has cytostatic, immunomodulator,  
CC neuroprotective, nootropic, anorectic, antidiabetic, antimicrobial, and

CC antilipaemic activity, and may have a use in gene therapy, and as a  
 CC vaccine. The polypeptides are encoded by NOVX polynucleotides comprising  
 CC any of the 303 fully defined nucleotide sequences given in the  
 CC specification. The polypeptide is useful in the manufacture of a  
 CC medicament for treating a syndrome associated with a human disease. The  
 CC polypeptide, polynucleotide and antibody are useful in diagnosing,  
 CC treating or preventing NOVX-associated disorders, e.g. cancer, cachexia,  
 CC Alzheimer's disease, Parkinson's disease, obesity, diabetes, infectious  
 CC diseases, metabolic syndrome X or dyslipidaemias. The nucleic acids are  
 CC further used as hybridisation probes, in chromosome mapping, tissue  
 CC typing, preventive medicine, and pharmacogenomics. The present sequence  
 CC represents a NOVX polypeptide of the invention.

XX

SQ Sequence 899 AA;

Query Match 97.9%; Score 4560.5; DB 8; Length 899;  
 Best Local Similarity 98.5%; Pred. No. 0;  
 Matches 862; Conservative 2; Mismatches 8; Indels 3; Gaps 3;

Qy	1	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	60
Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEWVCQVWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSSGEPTMEVRINVSRRQVEKVFGLLEEWVCQVWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDPSPDPNVYITREHS	180
Db	146	RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDPSPDPNVYITREHS	205
Qy	181	LVRQARLADTANYTCVAKNIVARRRSASAAVIVVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVRQARLADTANYTCVAKNIVARRRSASAAVIVVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDGWSWSPWSKWSACGLDCTHWSRECS	299
		:	
Db	266	SRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSVDGWSWSPWSKWSACGLDCTHWSRECS	325
Qy	300	PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIYVAVCLVLLLVLPILVYC	359
Db	326	PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIYVAVCLVLLLVLPILVYC	385
Qy	360	RKKEGLSDSDVADSSILTSQFQVSIKPSKADNPHLLTIQPDLS-TTTTYQGSICPRQDGP	419
Db	386	RKKEGLSDSDVADSSILTSQFQVSIKPSKADNPHLLTIQPDLS-TTTTYQGSICPRQDGP	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFN	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFN	504
Qy	480	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPP	564
Qy	540	GVLLTRPVILAMDHCGEPSDWSLRLKKQSCGSWE-DVLHLGEEAPSHLYYCQLEASA	598

Db	565	GVLTRPVILAMDHCGEPSDWSLRLKKQSCGSWEQDVLHLGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVAANKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	658
Db	625	CYVFTEQLGRFALVGEALSVAANKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	684
Qy	659	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRY	718
Db	685	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRY	744
Qy	719	LHCTFTLERVSPSTSDLACKLWWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	778
Db	745	LHCTFTLERVSPSTSDLACKLWWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	804
Qy	779	VGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLW	838
Db	805	VGPSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLW	864
Qy	839	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	873
Db	865	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899

Search completed: March 1, 2005, 08:56:45  
Job time : 202.724 secs

OM protein - protein search, using sw model

Run on: March 1, 2005, 08:42:47 ; Search time 52.3087 Seconds  
(without alignments)  
1245.848 Million cell updates/sec

Title: US-10-624-932-2\_COPY\_26\_898  
Perfect score: 4660  
Sequence: 1 QQSATVANPVPGANPDLLPH.....AVAGLGQPDAGLFTVSEAE 873

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result			% Query				Description
	No.	Score	Match	Length	ID		
	1	4511	96.8	898	2	US-08-808-982-5	Sequence 5, Appli
	2	4511	96.8	898	3	US-09-306-902A-5	Sequence 5, Appli
	3	2815.5	60.4	557	2	US-08-808-982-6	Sequence 6, Appli
	4	2815.5	60.4	557	3	US-09-306-902A-6	Sequence 6, Appli
	5	2560.5	54.9	943	2	US-08-808-982-7	Sequence 7, Appli
	6	2560.5	54.9	943	3	US-09-306-902A-7	Sequence 7, Appli
	7	2337.5	50.2	769	4	US-09-949-016-10665	Sequence 10665, A
	8	1881	40.4	886	4	US-09-969-532-16	Sequence 16, Appl
	9	1872.5	40.2	897	4	US-09-969-532-14	Sequence 14, Appl
	10	1864	40.0	900	4	US-09-969-532-12	Sequence 12, Appl
	11	1855.5	39.8	911	4	US-09-969-532-10	Sequence 10, Appl



12	1259.5	27.0	655	4	US-09-969-532-32	Sequence 32, Appl
13	1251	26.8	666	4	US-09-969-532-30	Sequence 30, Appl
14	1242.5	26.7	669	4	US-09-969-532-28	Sequence 28, Appl
15	1234	26.5	680	4	US-09-969-532-26	Sequence 26, Appl
16	1110	23.8	552	4	US-09-969-532-8	Sequence 8, Appli
17	1101.5	23.6	563	4	US-09-969-532-6	Sequence 6, Appli
18	1093	23.5	566	4	US-09-969-532-4	Sequence 4, Appli
19	1084.5	23.3	577	4	US-09-969-532-2	Sequence 2, Appli
20	488.5	10.5	321	4	US-09-969-532-24	Sequence 24, Appl
21	480	10.3	332	4	US-09-969-532-22	Sequence 22, Appl
22	471.5	10.1	335	4	US-09-969-532-20	Sequence 20, Appl
23	463	9.9	346	4	US-09-969-532-18	Sequence 18, Appl
24	296.5	6.4	1172	1	US-08-313-288B-19	Sequence 19, Appl
25	296.5	6.4	1172	4	US-09-949-016-6333	Sequence 6333, Ap
26	294	6.3	102	2	US-08-808-982-8	Sequence 8, Appli
27	294	6.3	102	3	US-09-306-902A-8	Sequence 8, Appli
28	276	5.9	954	4	US-09-854-845-14	Sequence 14, Appl
29	276	5.9	1049	4	US-09-854-845-2	Sequence 2, Appli
30	276	5.9	1093	4	US-09-854-845-4	Sequence 4, Appli
31	276	5.9	1151	4	US-09-854-845-10	Sequence 10, Appl
32	272.5	5.8	939	4	US-09-854-845-16	Sequence 16, Appl
33	272.5	5.8	1034	4	US-09-854-845-6	Sequence 6, Appli
34	272.5	5.8	1078	4	US-09-854-845-8	Sequence 8, Appli
35	272.5	5.8	1136	4	US-09-854-845-12	Sequence 12, Appl
36	272.5	5.8	1170	4	US-09-657-472-2	Sequence 2, Appli
37	268.5	5.8	239	5	PCT-US93-01652-1	Sequence 1, Appli
38	268.5	5.8	1170	1	US-08-313-288B-20	Sequence 20, Appl
39	249.5	5.4	441	3	US-08-985-526-3	Sequence 3, Appli
40	249	5.3	479	4	US-09-270-767-46823	Sequence 46823, A
41	243	5.2	469	1	US-08-313-288B-15	Sequence 15, Appl
42	243	5.2	484	4	US-09-949-016-9698	Sequence 9698, Ap
43	242.5	5.2	1045	4	US-09-949-016-11112	Sequence 11112, A
44	238	5.1	218	3	US-08-985-526-1	Sequence 1, Appli
45	226.5	4.9	1395	3	US-09-540-245A-15	Sequence 15, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-808-982-5

; Sequence 5, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsay

; APPLICANT: Masu, Masayuki

; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

```

;      ZIP: 94104
;      COMPUTER READABLE FORM:
;      MEDIUM TYPE: Floppy disk
;      COMPUTER: IBM PC compatible
;      OPERATING SYSTEM: PC-DOS/MS-DOS
;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/08/808,982
;      FILING DATE:
;      CLASSIFICATION: 530
;      ATTORNEY/AGENT INFORMATION:
;      NAME: OSMAN, RICHARD A
;      REGISTRATION NUMBER: 36,627
;      REFERENCE/DOCKET NUMBER: UC96-217
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415) 343-4341
;      TELEFAX: (415) 343-4342
;      INFORMATION FOR SEQ ID NO: 5:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 898 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: not relevant
;      TOPOLOGY: not relevant
;      MOLECULE TYPE: peptide
US-08-808-982-5

```

```

Query Match          96.8%; Score 4511; DB 2; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 838; Conservative 17; Mismatches 18; Indels 0; Gaps 0;

```

```

Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy      61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 120
        |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      86 DHVIERSTDSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 145

Qy     121 RLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146 YLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS 205

Qy     181 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
        |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     206 LVVRQARLADTANYTCVAKNIVARRRSTSAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

Qy     241 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWRSRECS DP 300
        |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     266 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSSWSKWSACGLDCTHWRSRECS DP 325

Qy     301 APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR 360
        ||||||||:| |||||||||:|:| |||||||:|:| |||| | |:|
Db     326 APRNGGEECRGADLDRNCTSDLCVHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCR 385

Qy     361 KKEGLDSDVADSSILTSQFQPVSIKPSKADNPHLLTIQPDLSSTTTTTYQGS LCP RQD GPS 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     386 KKEGLDSDVADSSILTSQFQPVSIKPSKADNPHLLTIQPDLSSTTTTTYQGS LCP RQD GPS 445

```

Qy	421	PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNF	480
		:             :	
Db	446	PKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMAYGTNF	505
Qy	481	LGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG	540
		:	
Db	506	LGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPVVSCGPPG	565
Qy	541	VLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACY	600
		:	
Db	566	VLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWEDVLHLGEESP SHLYYCQLEAGACY	625
Qy	601	VFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLH DTHDALKEVVQLE	660
		:	
Db	626	VFTEQLGRFALVGEALSVA ATKRLRLLLFAPVACTSLEYNIRVYCLH DTHDALKEVVQLE	685
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRLSIH DVPSSLWKS KLLVSYQEIPFYHIWNGTQRYLH	720
		:	
Db	686	KQLGGQLIQEPRVLHFKDSYHNLRLSIH DVPSSLWKS KLLVSYQEIPFYHIWNGTQQYLH	745
Qy	721	CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	780
		:        :      :	
Db	746	CTFTLERINASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEGGVPALVG	805
Qy	781	PSAFKIPFLIRQKIISLDPCCRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA	840
		:	
Db	806	PSAFKIPFLIRQKIIASLDPCCSRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA	865
Qy	841	RHFPGNGLSQLAAAVAGLGQPDAGLFTVSEAE C	873
Db	866	RHFPGNGLGQLAAAVAGLGQPDAGLFTVSEAE C	898

## RESULT 2

US-09-306-902A-5

; Sequence 5, Application US/09306902A

; Patent No. 6277585

### ; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

```

;      SOFTWARE: PatentIn Release #1.0, Version #1.30
;
;      CURRENT APPLICATION DATA:
;      APPLICATION NUMBER: US/09/306,902A
;      FILING DATE: 07-May-1999
;      CLASSIFICATION: <Unknown>
;
;      ATTORNEY/AGENT INFORMATION:
;      NAME: OSMAN, RICHARD A
;      REGISTRATION NUMBER: 36,627
;      REFERENCE/DOCKET NUMBER: UC96-217
;
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415) 343-4341
;      TELEFAX: (415) 343-4342
;
;      INFORMATION FOR SEQ ID NO: 5:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 898 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: not relevant
;      TOPOLOGY: not relevant
;
;      MOLECULE TYPE: peptide
;      SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-306-902A-5

```

```

Query Match          96.8%; Score 4511; DB 3; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 838; Conservative 17; Mismatches 18; Indels 0; Gaps 0;

```

```

Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy      61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 120
        |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      86 DHVIERSTDSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 145

Qy     121 RLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146 YLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS 205

Qy     181 LVVRQARLADTANYTCVAKNIVARRRSASAIVIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
        |||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     206 LVVRQARLADTANYTCVAKNIVARRRSTSAIVIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

Qy     241 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSWSPWSKWSACGLDCTHWSRECS DP 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     266 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSWSWSPWSKWSACGLDCTHWSRECS DP 325

Qy     301 APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIYVAVCLVLLLLVLILVYCR 360
        |||||:| |||||:|:| |||||:|:| ||||| ||| | :||
Db     326 APRNGGEECRGADLDRNCTSDLCVHSASGPEDVALYVGLIYVAVCLVLLLLVLILVYCR 385

Qy     361 KKEGLDSDVADSSILTSFGFQPVSIKPSKADNPHLLTIQPDLSSTTTTTYQGSLCPRQDGPS 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     386 KKEGLDSDVADSSILTSFGFQPVSIKPSKADNPHLLTIQPDLSSTTTTTYQGSLCSRQDGPS 445

Qy     421 PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMAYGTGTFNF 480
        ||||:||||||| |||||:|||||||:||||||| ||||| |||||
Db     446 PKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMAYGTGTFNF 505

```

Qy 481 LGGRLMIPNTGISLLIPPDaiprgkiyeiyltlhkpedvrlplagcqtllspivscgppg 540  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||||||  
 Db 506 LGGRLMIPNTGISLLIPPDaiprgkiyeiyltlhkpedvrlplagcqtllspvscgppg 565  
 Qy 541 VLLTRPVILAMDHCGEPSPDSWSLRLKKQSCGWSWEDVLHLGEEAPSHLYYCQLEASACY 600  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:||||||| |||  
 Db 566 VLLTRPVILAMDHCGEPSPDSWSLRLKKQSCGWSWEDVLHLGEEAPSHLYYCQLEAGACY 625  
 Qy 601 VFTEQLGRFALVGEALSVAaakrlklllfapvactsleyNIRVYCLHDTHDALKEVVQLE 660  
 ||||||||||||||||||||| |||:|||||||||||||||||||||||||||||||||||  
 Db 626 VFTEQLGRFALVGEALSVAATKRLRlllfapvactsleyNIRVYCLHDTHDALKEVVQLE 685  
 Qy 661 KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLH 720  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||:|||  
 Db 686 KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQQYLH 745  
 Qy 721 CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG 780  
 |||||||:| |||||||:|||||||||:||||||| |||||||  
 Db 746 CTFTLERINASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEGGVPALVG 805  
 Qy 781 PSAFKIPFLIRQKIISLDPCCRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEA 840  
 |||||||||||||||:||||| |||||||||||||||||||||||||||||||||||||  
 Db 806 PSAFKIPFLIRQKIIASLDPCCRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEA 865  
 Qy 841 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEc 873  
 ||||||| |||||||||||||||||||||  
 Db 866 RHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAEc 898

RESULT 3

US-08-808-982-6

; Sequence 6, Application US/08808982

; Patent No. 5939271

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; APPLICANT: Leonardo, E. David

; APPLICANT: Hink, Lindsay

; APPLICANT: Masu, Masayuki

; APPLICANT: Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/808,982

; FILING DATE:

```

;      CLASSIFICATION: 530
;      ATTORNEY/AGENT INFORMATION:
;      NAME: OSMAN, RICHARD A
;      REGISTRATION NUMBER: 36,627
;      REFERENCE/DOCKET NUMBER: UC96-217
;      TELECOMMUNICATION INFORMATION:
;      TELEPHONE: (415) 343-4341
;      TELEFAX: (415) 343-4342
;      INFORMATION FOR SEQ ID NO: 6:
;      SEQUENCE CHARACTERISTICS:
;      LENGTH: 557 amino acids
;      TYPE: amino acid
;      STRANDEDNESS: not relevant
;      TOPOLOGY: not relevant
;      MOLECULE TYPE: peptide
US-08-808-982-6

```

```

Query Match          60.4%; Score 2815.5; DB 2; Length 557;
Best Local Similarity 96.8%; Pred. No. 3.5e-258;
Matches 539; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

```

```

Qy      318 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSIITS 377
        ||||| ||:|||||||||||||||||||||||||||||||||||||||||
Db       1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSIITS 60

Qy      378 GFQPVSIKPSKADNPHELLTIQPDLS'TTTT'YQGS'LCPRQDGSPK'FQLTNGHLLSPLGGG 437
        |||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GFQPVSIKPSKADNPHELLTIQPDLS'TTTT'YQGS'LCPRQDGSPK'FQLTNGHLLSPLGGG 120

Qy      438 RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGT'FNFLGGRLMIPNTGISLLIP 497
        |||||||||||||||||||||||||||||||||||||||||||||||
Db      121 RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGT'FNFLGGRLMIPNTGISLLIP 180

Qy      498 PDAIPRGKIYEIYLT'LHKPEDVRLPLAGCQTLLS'PIVSCGPPGVLLTRPVILAMDHC'GEP 557
        |||||||||||||||||||||||||||||||||||||||||||||||
Db      181 PDAIPRGKIYEIYLT'LHKPEDVRLPLAGCQTLLS'PIVSCGPPGVLLTRPVILAMDHC'GEP 240

Qy      558 SPDSWSLRLKKQSC'EGSWEDVLHLGEEAPSHLYYC'QLEASACYVFTEQLGRFALVGEALS 617
        ||||| |||||||||||||||||||||||||||||||||||||||||||
Db      241 SPDSWSLALKKQSC'EGSWEDVLHLGEEAPSHLYYC'QLEASACYVFTEQLGRFALVGEALS 300

Qy      618 VAAAKRLKLLL'FAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK 677
        |||||||||||||||||||||||||||||||||||||||||||||||
Db      301 VAAAKRLKLLL'FAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHLX 360

Qy      678 DSYHNLRLSIH'DVPSSLWKS'KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLAC 737
        ||||| || |||||||||||||||||||||||||||||||||||||||
Db      361 DSYHNLXLSXH'DVPSSLWKS'KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLAC 420

Qy      738 KLWVWQVEGDGQ'SFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKI'ISS 797
        |||||||||||||||||||||||||||||||||||||||||||||
Db      421 KLWVWQVEGDGQ'SFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKI'ISS 480

Qy      798 LDPPCRRGADWRTLAQKLHLD'SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAG 857
        |||||||||||||||||||||||||||||||||||||||||||||
Db      481 LDPPCRRGADWRTLAQKLHLD'SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAG 540

```

Qy 858 LGQPDAGLFT-VSEAEC 873  
| : |||||  
Db 541 TXPAGRWLLSQCSEAEC 557

RESULT 4

US-09-306-902A-6

; Sequence 6, Application US/09306902A  
; Patent No. 6277585  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; Leonardo, E. David  
; Hink, Lindsay  
; Masu, Masayuki  
; Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/306,902A  
; FILING DATE: 07-May-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 557 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-306-902A-6

Query Match 60.4%; Score 2815.5; DB 3; Length 557;  
Best Local Similarity 96.8%; Pred. No. 3.5e-258;  
Matches 539; Conservative 2; Mismatches 15; Indels 1; Gaps 1;

Qy 318 NCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 377  
||||| ||:||||||||||||||||||||||||||||||||||||||  
Db 1 NCTSDLXVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILTS 60





```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/808,982
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 943 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-808-982-7

```

```

Query Match          54.9%; Score 2560.5; DB 2; Length 943;
Best Local Similarity 54.0%; Pred. No. 1.4e-233;
Matches 496; Conservative 139; Mismatches 215; Indels 69; Gaps 14;

```

```

Qy      11 PGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG 70
      | | : |||||:|||| ||||| | | :| |||||:||||||| | || : | |
Db      38 PSAPAEQLPHFLLEPEDAYIVKNKPVELHCRAFPATQIYFKCNGEWVSQKGHVTVQESLDE 97

Qy      71 SSSLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP 130
      ::|| ||:| |||||:||||:||||||| |||||:||||| ||||:|||
Db      98 ATGLRIREVQIEVSRQQVEELFGLLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEP 157

Qy     131 LAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHSLVVRQARLAD 190
      |||| | : ::| |||||:| |||||:||||:| | : | :|:|:||||:|
Db     158 LAKEVPLDHEVLLQCRPPEGVPVAEVEWLKNEDVIDPAQDTNFLTIDHNLIRQARLSD 217

Qy     191 TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPL 250
      ||||| |||||:||| :| ||||| ||||:| ||| || |||||:|:|||||
Db     218 TANYTCVAKNIVAKRRSTTATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRCTNPAPL 277

Qy     251 NGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWRSRSCDPAPRNGGEECQ 310
      ||||| |||| | :|||:|: ||||| :| ||||| | | :||| :|
Db     278 NGGAFCEGQACQKTACTTVCPVDGAWTEWSKWSACSTECAHWRSRCEMAPPQNGGRDCS 337

Qy     311 GTDLDTNRCTSDLCV---HSASGPE-----DVALYVGL-IAVAVCLVLLLLVLILVY 358
      || ||:| || || : : | : |||| || :|| | | :| :| :||
Db     338 GTLLDSKNCTDGLCVLNQRTLNPKSRPLEPSGDVALYAGLVAVFVVLAVLMAVGIVY 397

Qy     359 CRKKEGLDSDVADSS-ILTSGFQPVSIKPSKADNPHLL--TIQPDLSSTTTTYYQGSCLPR 415
      | :|: || | || || ||: | : | || | : ||| : | :| :
Db     398 RRNCRDFDITDSSAALTGGFHPVNFKTARPSNPQLLHPSAPPDLTASAGIYRGPVYAL 457

```

Qy 416 QDGPSPKFQLTNGHLLSPL-----GGG----- 437  
 Db 458 QDS-ADKIPMTNSPLLDPLPSLKIKVYDSSTIGSGAGLADGADLLGVLPPTYPGDFSRD 516

Qy 438 RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNFGGRLMIPNTGISLLIP 497  
 Db 517 THFLHLRS-----ASLGSQ-HLLGLPRDPSSSVSGTFGCLGGRLTIPGTGVSLVLP 566

Qy 498 PDAIPRGKIYEIYLT LHKPEDVRLPLA-GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE 556  
 Db 567 NGAI PQGKFYDLYLRINKTEST-LPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPHCAE 625

Qy 557 PSPDSWSRLKKQSCGWSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 616  
 Db 626 VIAGDWIFQLKTQAHQGHWEVVTLDEETLNTPCYQCQLEAKSCHILLDQLGTYVFTGESY 685

Qy 617 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 676  
 Db 686 SRSVAVKRLQLAIFAPALCTSLEYSLRVYCLEDTPAALKEVLELERTLGGYLVEEPKTLLE 745

Qy 677 KDSYHNLRLSIHDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTDLA 736  
 Db 746 KDSYHNLRLSLHDIPHAHWSKLLAKYQEIPFYHVWNGSQKALHCTFTLERHSLASTEFT 805

Qy 737 CKLWWVQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL--VGPSAFKIPFLIRQKI 794  
 Db 806 CKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDALCSAPGNAATTQLGPYAFKIPLSIRQKI 864

Qy 795 ISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAA 854  
 Db 865 CNSLDAPNSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLWEARQQDDGDLNSLASA 924

Qy 855 VAGLGQPDAGLFTVSEAE 873  
 Db 925 LEEMGKSEMLVAMTTDGD 943

RESULT 6

US-09-306-902A-7

; Sequence 7, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

```

;           MEDIUM TYPE: Floppy disk
;           COMPUTER: IBM PC compatible
;           OPERATING SYSTEM: PC-DOS/MS-DOS
;           SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
;           APPLICATION NUMBER: US/09/306,902A
;           FILING DATE: 07-May-1999
;           CLASSIFICATION: <Unknown>
;
; ATTORNEY/AGENT INFORMATION:
;           NAME: OSMAN, RICHARD A
;           REGISTRATION NUMBER: 36,627
;           REFERENCE/DOCKET NUMBER: UC96-217
;
; TELECOMMUNICATION INFORMATION:
;           TELEPHONE: (415) 343-4341
;           TELEFAX: (415) 343-4342
;
; INFORMATION FOR SEQ ID NO: 7:
;           SEQUENCE CHARACTERISTICS:
;             LENGTH: 943 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: not relevant
;             TOPOLOGY: not relevant
;
;           MOLECULE TYPE: peptide
;           SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-306-902A-7

```

```

Query Match          54.9%; Score 2560.5; DB 3; Length 943;
Best Local Similarity 54.0%; Pred. No. 1.4e-233;
Matches 496; Conservative 139; Mismatches 215; Indels 69; Gaps 14;

```

```

Qy      11 PGANPDLLPHFLVEPEDVYIVKNKPVLVCKAVPATQIFFKNGEWVRQVDHVIERSTDG 70
      | | : ||||:|||| ||||| | | :| ||||:|||||| | || : | |
Db      38 PSAPAEQLPHFLLEPEDAYIVKNKPVELHCRAFPATQIYFKNGEWVSQKGHVTQESLDE 97

Qy      71 SSSLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP 130
      :|| | ||:| |||||:||||:||||||| |||||:||||| ||||:| |
Db      98 ATGLRIREVQIEVSRQQVEELFGLLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEP 157

Qy     131 LAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDNVYITREHSLVVRQARLAD 190
      |||| | : :| |||||:| |||||:||||:| | :| :|:|:| ||||:|
Db     158 LAKEVPLDHEVLLQCRPPEGVPVAEVEWLKNEDVIDPAQDTNFLTIDHNLIIRQARLSD 217

Qy     191 TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPL 250
      ||||| |||||:| | :| ||||| ||||:| || | | |||||:|:| |||||
Db     218 TANYTCVAKNIVAKRRSTTATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTCTCTNPAPL 277

Qy     251 NGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQ 310
      ||||| |||| | :|||:|: ||||| :| ||||| | | :||| :|
Db     278 NGGAFCEGQACQKTACTTVCPVDGAWTEWSKWSACSTECAHWSRECMAPPPQNGGRDCS 337

Qy     311 GTDLDTRNCTSDLCV---HSASGPE-----DVALYVGL-IAVAVCLVLLLLVLILVY 358
      || ||:| | || : : | : |||| || :| | | :| :| :| |
Db     338 GTLLDSKNCTDGLCVLNQRTLNDPKSRPLEPSGDVALYAGLVVAVFVVLAVLMAVGVIYV 397

Qy     359 CRKKEGLDSDVADSS-ILTSGFQPVSIKPSKADNPHELL--TIQPDLTSTTTTTYQGSCLCPR 415
      | :|: || | || || ||: | :| || | : ||| : | :| :
Db     398 RRNCRDFDITDSSAALTGGFHPVNFKTARPSNPQLLHPSAPPDLTASAGIYRGPVYAL 457

```

Qy 416 QDGPSPKFQLTNGHLLSPL-----GGG----- 437  
 || : | :|| || || | |  
 Db 458 QDS-ADKIPMTNSPLLDPLPSLKIKVYDSSTIGSGAGLADGADLLGVLP PGTYPGDFSRD 516  
 Qy 438 RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNFNLGGRLMIPNTGISLLIP 497  
 | || | : | : | : || | : || | |||| | | ||: ||: |  
 Db 517 THFLHLRS-----ASLSQ-HLLGLPRDPSSSVSGTFGCLGGRLTIPGTGVSLLPV 566  
 Qy 498 PDAIPRGKIYEIYLTLLHKPEDVRLPLA-GCQTLSPIVSCGPPGVLLTRPVILAMDHCGE 556  
 |||: || |: || :| | ||: | ||: || |: || |: || |: || : || |  
 Db 567 NGAI PQGKFYDLYLRINKTEST-LPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVP HCAE 625  
 Qy 557 PSPDSWSLRLKKQSCGWSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 616  
 | : || |: | ||: | || : |||| : ||: : || : ||:  
 Db 626 VIAGDWIFQLKTQAHQGHWEVVTLD EETLNTPCYCQLEAKSCHILLDQLGTYVFTGESY 685  
 Qy 617 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 676  
 | : | ||: | : || | ||||: |||| | | ||||: ||: || | ||: | |  
 Db 686 SRS AVKRLQLAIFAPALCTSLEYSLRVYCLEDTPAALKEVLELERTLGGYLVEEPKTL LF 745  
 Qy 677 KDSYHNLRSLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 736  
 ||||| ||||: ||: | : |||| | ||||: ||: | : ||||| | : ||:  
 Db 746 KDSYHNLRSLSLHDIPHAHWSKLLAKYQEIPFYHVWNGSQKALHCTFTLERHSLASTEFT 805  
 Qy 737 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL--VGPSAFKIPFLIRQKI 794  
 ||: | ||||: || | : : : | | | | | | : || |||| | ||||  
 Db 806 CKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDALCSAPGNAATTQLGPYAFKIPLSIRQKI 864  
 Qy 795 ISSLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAA 854  
 : || | | || || |||| : | : ||: ||: | || : ||: |||| : ||: | ||:  
 Db 865 CNSLDAPNSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLEARQDDGDLNSLASA 924  
 Qy 855 VAGLGQPDAGLFTVSEAEC 873  
 : : || : : : : |  
 Db 925 LEEMGKSEMLVAMTTDGDC 943

# RESULT 7

US-09-949-016-10665

; Sequence 10665, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSEQ for Windows Version 4.0

```
; SEQ ID NO 10665
;   LENGTH: 769
;   TYPE: PRT
;   ORGANISM: Human
US-09-949-016-10665
```

Query Match 50.2%; Score 2337.5; DB 4; Length 769;  
Best Local Similarity 56.1%; Pred. No. 1.5e-212;  
Matches 432; Conservative 136; Mismatches 179; Indels 23; Gaps 7;

Qy	122	LRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSL	181
Db	3	LRKTFEQEPLGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEDIIDPVEDRNFYITIDHNL	62
Qy	182	VVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRS	241
Db	63	IIKQARLSDTANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRT	122
Qy	242	RSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DPA	301
Db	123	RTCTNPAPLNGGAFCEGQSVQKIACTTLC PVDGRWTPWSKWSTCGTECTHWRRRECTAPA	182
Qy	302	PRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCR	360
Db	183	PKNGGKDCDGLVLQSKNCTDGLCMQTAPDSDVALYVGIVIAVIVCLAISVVVALFVYRK	242
Qy	361	KKEGLDSDVDASSILTSGFQPVSIKPSKADNPHELLTIQPDLSTTTTTTYQGS LCPRODGPS	420
Db	243	NHRDFESDIIDSSALNGGFQPVNIKAARQD---LLAVPPDLTSAAAMYRGVPYALHD-VS	298
Qy	421	PKFQLTNGHLLSPLGGGRHTLHHSS----PTSEAEFEVSRLS---TQNYF-----R	464
Db	299	DKIPMTNSPILDPLPNLKIKVYNTSGAVTPQDDLSEFTSKLSPQMTQS LLENEALS LKNQ	358
Qy	465	SLPRGT--SNMTYGTFFNLGGRLMIPNTGISLLIPPD AIPRGKIYEIYLT LHKPEDVRLP	522
Db	359	SLARQTDPSCTAFGSFNSLGGHLIVPNSGVSL LI PAGAI PQGRVYEMYVTVHRKETMRPP	418
Qy	523	LAGCQTLTSPIVSCGPPGVLLTRPVILAMDHCGEPSPD SWSLRLKKQSCEGSWEDVLHLG	582
Db	419	MDDSQTLTTPVVSCGPPGALLTRPVVLTMHHCADPN TEDWKILLKNQAAQGQWEDVVVVG	478
Qy	583	EEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLL FAPVACTSLEYNIR	642
Db	479	EENFTTPCYIQLDAEACHILTENLSTYALVGHSTTKAA AKRLKLAI FGPLCCSSLEYSIR	538
Qy	643	VYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSY HNLRLSIHDVPSSLWKS KLLVS	702
Db	539	VYCLDDTQDALKEILHLERQMGGQLLEEPKALHFKG STHNLRLSIHDIAHSLWKS KLLAK	598
Qy	703	YQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKL WVWQVEGDGQSFSINFNITKDTR	762
Db	599	YQEIPFYHVWSGSQRN LHCTFTLERFSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEPT	658
Qy	763	FAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLD PPCCRAGADWRTLAQKLHLD SHLS	822
Db	659	GIDLPLLD PANTITTVTGPSAFSIP LPIROKLCSSLDAPOTRGHDWRMLAHKLNLD RYLN	718

Qy 823 FFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 872  
 :||:| ||| :||:||||:|||| ||| : :|: : : :| :  
 Db 719 YFATKSSPTGVILDLWEAQNFPDGNLSMLAAVLEEMGRHETVVSIAAEGQ 768

RESULT 8

US-09-969-532-16

; Sequence 16, Application US/09969532

; Patent No. 6777232

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides  
 Encoding the Same

; FILE REFERENCE: LEX-0244-USA

; CURRENT APPLICATION NUMBER: US/09/969,532

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US 60/237,280

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 886

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-969-532-16

Query Match 40.4%; Score 1881; DB 4; Length 886;

Best Local Similarity 41.9%; Pred. No. 4.3e-169;

Matches 376; Conservative 155; Mismatches 261; Indels 106; Gaps 12;

Qy 10 VPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTD 69  
 :| | | |||: ||:| ||:| :| | || | ||||| ||| :|| | :|  
 Db 46 IPSA-PGTLPHFIEEPDDAYIISKNPIALRCKARPAMQIFFKCNGEWVHQNEHVSEETLD 104  
 Qy 70 GSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQE 129  
 ||| | |||:|||| | |:||||| ||:|:| | :|| |||||:  
 Db 105 ESSGLKVREVFINVTRQQVEDFHGPEDYWCQCVAWSHLGTSKSRKASVRIAYLRKNFEQD 164  
 Qy 130 PLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLA 189  
 | :|| :| || | |||||:| |||||:|: :| | |: :|:|:| ||||:  
 Db 165 PQGREVPPIEGMIVLHCRPPEGVPAAEVEWLKNEEPIDSEQDENIDTRADHNLIIHQARLS 224  
 Qy 190 DTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAP 249  
 |: ||||:| ||||:| || |:|||:| | |:|||||  
 Db 225 DSGNYTCMAANIVAKRRSLSATVVVYVDGSWEVWSEWSVCSP----- 266  
 Qy 250 LNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEEC 309  
 :| | | |||: | ||||: |  
 Db 267 -----ECEHLRIRECTAPPPRNGGKFC 288  
 Qy 310 QGTDLDTRNCTSDLCVHSASGPEdVALYVGLIAVAVCLVLLLVLILVYCRKKEGLDSDV 369  
 :| :| ||| ||: |:|| || || | :|: :| | : ||  
 Db 289 EGLSQESENCTDGLCILGIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDYGVDV 347  
 Qy 370 ADSSILTSGFQPVSIKPSKADNPHELL--TIQPDLSSTTTTTYQGSCLPRQDGPSPKFQLTN 427









Db 267 -----ECEHLRIRECTAPPPRNGGKFC 288  
 Qy 310 QGTDLDTNRCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDV 369  
 :| :: ||| ||: |:||| || || | :||: : :| | : ||  
 Db 289 EGLSQESENCTDGLCILGIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDYGV DV 347  
 Qy 370 ADSSILTSQFQPVSIKPSKA-----DNPHLL--TIQPDLS TTTT TTYQGS LC 413  
 ||| || ||| : | : | || : ||| | : || | : |  
 Db 348 IDSSALTGGFQTFNFKTVRQAKNIMELMIQEKSFGNSLLLNSAMQPD L-TVSR TYSGPIC 406  
 Qy 414 PRQDGPSPKFQLTNGHLLSPLG-----GGRHT 440  
 || | | : | | : || | | :  
 Db 407 -LQD-PLDKELMTESL FNPLSDIKVKVQSSFMVSLGVSERAEYHGKNHSRTFPHGNNHS 464  
 Qy 441 LHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGT FNFLGGRLMI PNTGISLLIPPDA 500  
 | :: :: || || | | | | ||||: |||: |||| |  
 Db 465 FSTMHPRNKM-PYIQNLS-----SLPTRTELRTTG VFGHLGGRLVMPNTGV SLLI PHGA 517  
 Qy 501 IPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPD 560  
 || :|||::: | | | : ||| | :||| :||| :||| :||| :  
 Db 518 IPEENSWEIYMSINQGE P-SLQSDGSEVLLSPEVTCGPPDMI VTT P FALTIPHCADVSSE 576  
 Qy 561 SWSLRLKKQSCGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA A 620  
 |:: |||:: : | ||:| : :|: | || | : ||:| : | :|| || : : |  
 Db 577 HWNHLLKKRTQQGKWEVMSVEDESTS--CYCLDPFACHVLLDSFGTYALTGEPI TDCA 634  
 Qy 621 AKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSY 680  
 |:||: :| ::| ||:| |||: :| | :||| | : ||||: |||: |||| :  
 Db 635 VKQLKVAVFGCMSCNSLDYNLRVYCV DNTPCAFQEVVSDERHQGGQLLEPKLLHFKGNT 694  
 Qy 681 HNLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLW 740  
 :||:| :| | | : | : ||:| :| :: || | :||| :|:| :|:| :  
 Db 695 FSLQISVLDIPPFLWRIPFTACQEV PFSRVWC SNRQPLHCAFS LERYTPTTTQLSCKIC 754  
 Qy 741 VWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDP 800  
 : ||:| | : : : | : : || || ||||| :||:| ||:| :  
 Db 755 IRQLKGHEQILQVQTSILESERETITFFAQEDSTFPAQTGPKAFKIPYSIRQRICATFDT 814  
 Qy 801 PCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQ 860  
 | :| ||: |||| :: :||:| :| ||:| ||||| :||:| || | : :|  
 Db 815 PNAKGKD WQMLAQKNSINRNLSYFATQSSPSAVILNLWEARHQHDGDLDSLACALEEIGR 874  
 Qy 861 PDAGLFTVSEAE 872  
 | :||:  
 Db 875 THTKLSNISESQ 886

RESULT 11

US-09-969-532-10

; Sequence 10, Application US/09969532

; Patent No. 6777232

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides  
Encoding the Same

; FILE REFERENCE: LEX-0244-USA



```

      : || : | : |:: ||:: :| ||:: : :|: | || : ||:| : | :
Db      577 TIPHCADVSEHWNHLLKKRTQQGKWEEVMSVEDESTS--CYCLDPPFACHVLLDSFGTY 634

Qy      610 ALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQ 669
      || || :: | |::|: :| ::| ||::||::||:: :| | :|| | : ||:::
Db      635 ALTGEPITDCAVKQLKVAVFGCMSCNSLDYNLRVYCDNTPCAFQEVVSDERHQGGQLE 694

Qy      670 EPRVLHFKDSYHNLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVS 729
      ||::||| : :|::|: |:: ||: | : ||::| :| :: || | :||| :
Db      695 EPKLLHFKGNTFSLQISVLDIPFLWRIKPFTACQEVVPSRVWCSNRQPLHCAFSLERYT 754

Qy      730 PSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFL 789
      |::| :||: : |::| | : :| : | : :: || || ||||:
Db      755 PTTTQLSCKICIRQLKGHEQILQVQTSILESERETITFFAQEDSTFPAQTGPKAFKIPYS 814

Qy      790 IRQKIISLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLS 849
      ||::| :: | | :| ||: ||| :: :||:|:: ||::| ||||| :|:|
Db      815 IRQRICATFDT PNAKGKDWQMLAQKNSINRNLSYFATQSSPSAVILNLWEARHQHDGDL 874

Qy      850 QLAAAVAGLGQPDAGLFTVSEAE 872
      || | : :|: | :||::
Db      875 SLACALEEIGRTHTKLSNISESQ 897

```

RESULT 12

US-09-969-532-32

; Sequence 32, Application US/09969532

; Patent No. 6777232

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides Encoding the Same

; FILE REFERENCE: LEX-0244-USA

; CURRENT APPLICATION NUMBER: US/09/969,532

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US 60/237,280

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 32

; LENGTH: 655

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-969-532-32

Query Match 27.0%; Score 1259.5; DB 4; Length 655;

Best Local Similarity 36.8%; Pred. No. 2.7e-110;

Matches 262; Conservative 128; Mismatches 216; Indels 105; Gaps 11;

```

Qy      197 VAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPLNGGAFC 256
      :| ||||:| | ||:|:| | | :||| |
Db      1 MAANIVAKRRSLSATVVVYVDGSWEVWSEWSVCSP----- 35

Qy      257 EGQNVQKTACATLCPVDGSWPWSKWSACGLDCTHWSRECSDPAPRNGGEECQGTDLDT 316
      :| | | ||: | ||||: |:| ::
Db      36 -----ECEHLRIRECTAPPPRNGGKFCEGLSQES 64

```

Qy 317 RNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVDADSSILT 376  
 ||| ||: |:||| || | | :||: : :| | : || ||| ||  
 Db 65 ENCTDGLCILGIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDYGVDVIDSSALT 123

Qy 377 SGFQPVSIKPSKADNPHLL--TIQPDLSSTTTTTYQGSLCPRQDGPSPKFQLTNGHLLSPL 434  
 ||| : | : | || :||| | : || | :| || | :| | :||  
 Db 124 GGFQTFNFKTVRQGNLNSAMQPD-LTVSRTYSGPIC-LQD-PLDKELMTESSLFNPL 180

Qy 435 G-----GGRHTLHHSSPTSEAEFVSRLSTQN 461  
 | | : | : : ||  
 Db 181 SDIKVKVQSSFMVSLGVSERAEYHGKNHSRTFPHGNNHSFSTMHPRNKM-PYIQNLS--- 236

Qy 462 YFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRL 521  
 ||| | | | ||||:||||:|||| ||| :|||::: | |  
 Db 237 ---SLPTRTELRTTGVFHGLGGRLVMPNTGVSLIPHGAIPPEENSWEIYMSINQGEPSL 292

Qy 522 PLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHL 581  
 | : |||| |:||| :|| | | : | : | : |||: | |||: | :  
 Db 293 QSDGSEVLLSPEVTCGPPDMIVTTPFALTIPHCADVSSEHWNHLKKRTQQGKWEVMSV 352

Qy 582 GEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAABKRLKLLLFAPVACTSLEYNI 641  
 :|: | || | : ||: | : | || || : : | |:||: | :| ||:|:  
 Db 353 EDESTS--CYCLDPFACHVLLDSFGTYALTGEPIIDCAVKQLKVAVFGCMSCNSLDYNL 410

Qy 642 RVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSCLLV 701  
 ||||: :| | :|| | : ||||:||||:|||| : :||: | :| ||: |  
 Db 411 RVYCVDNTPCAFQEVVSDERHQGGQLLEPKLLHFKGNTFSLQISVLDIPFLWRKIPFT 470

Qy 702 SYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDT 761  
 : ||:| | : : || |:|| :||: |:||: | :|| | : :| :  
 Db 471 ACQEVFPFSRVWCNSNRQPLHCAFSLERYTPTTTQLSCKICIRQLKGHEQILQVQTSILESE 530

Qy 762 RFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDLHSHL 821  
 | : : || || ||||: |||: | : | : ||| ||| : :|  
 Db 531 RETITFFAQEDSTFPAQTGPKAFKIPYSIRQRICATFDTFNAKGKDWQMLAQKNSINRNL 590

Qy 822 SFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 872  
 |:||: ||:|:||||| :|: | | | : :| | :||:  
 Db 591 SYFATQSSPSAVILNLWEARHQHDGDLDSLACALEEIGRTHTKLSNISESQ 641

# RESULT 13

US-09-969-532-30

; Sequence 30, Application US/09969532

; Patent No. 6777232

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides  
 Encoding the Same

; FILE REFERENCE: LEX-0244-USA

; CURRENT APPLICATION NUMBER: US/09/969,532

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US 60/237,280

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 33



Db 591 LAQKNSINRNLSYFATQSSPSAVILNLWEARHQHDGDLDSLACALEEIGRTHTKLSNISE 650  
Qy 871 AE 872  
::  
Db 651 SQ 652

RESULT 14

US-09-969-532-28

; Sequence 28, Application US/09969532

; Patent No. 6777232

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides  
Encoding the Same

; FILE REFERENCE: LEX-0244-USA

; CURRENT APPLICATION NUMBER: US/09/969,532

; CURRENT FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: US 60/237,280

; PRIOR FILING DATE: 2000-10-02

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 28

; LENGTH: 669

; TYPE: PRT

; ORGANISM: homo sapiens

US-09-969-532-28

Query Match 26.7%; Score 1242.5; DB 4; Length 669;

Best Local Similarity 36.1%; Pred. No. 1.2e-108;

Matches 262; Conservative 128; Mismatches 216; Indels 119; Gaps 12;

Qy 197 VAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPLNGGAFC 256  
:| ||||:|||| || |:||||:| | |:|||||  
Db 1 MAANIVAKRRSLSATVVVYVDGSWEVWSEWSVCSP----- 35  
Qy 257 EGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWRSRECSDPAPRNGGEECQGTDLDT 316  
:| | | |||: | ||||: |:| ::  
Db 36 -----ECEHLRIRECTAPPPRNGGKFCEGLSQES 64  
Qy 317 RNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLDSDVADSSILT 376  
||| ||: |:||| || || |:|::: |:| | | ||| ||  
Db 65 ENCTDGLCILGIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDYGVDVIDSSALT 123  
Qy 377 SGFQPVSISKPSKA-----DNPHELL--TIQPDLTSTTTTYQGSICPRQDGPS 420  
||| :| : | || :||| | :| | | :| |||  
Db 124 GGFQTFNFKTVRQAKNIMELMIQEKSFGNSLLNSAMQPDL-TVSRTYSGPIC-LQD-PL 180  
Qy 421 PKFQLTNGHLLSPLG-----GGRHTLHHSSPT 447  
| :| | :|| | | :| |  
Db 181 DKELMTESSLFNPLSDIKVKVQSSFMVSLGVSERAEYHGKNHSRTFPHGNHHSFSTMHPR 240  
Qy 448 SEAEFVSRLSTQNYFRSLPRGTSNMTYGTENFLGGRLMIPNTGISLLIPPDAIPRGKIY 507  
:: :: || || | | | ||||:||||:|||| ||| :  
Db 241 NKM-PYIQNLS-----SLPTRTELRTTGVEGHLGGRLVMPNTGVSLIPHGAIPPENSW 293

Qy 508 EIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSWSLRK 567  
 |||::: | | : ||| |::| | : | : | : | : ||  
 Db 294 EIYMSINQGEP-SLQSDGSEVLLSPEVTCGPPDMIVTTPFALTIPHCADVSEHWNHLK 352

Qy 568 KQSCGSEWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLL 627  
 |:: :| ||:: : :|: | | : ||: | : | : || | : : | |::| :  
 Db 353 KRTQQGKWEVMSVEDESTS--CYCLLDPFACHVLLDSFGTYALTGEPI TDCAVKQLKVA 410

Qy 628 LFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSI 687  
 :| ::| ||::|:||||: :| | :||| |: |||::|:|:|:|:| : :|::|:  
 Db 411 VFGCMSCNSLDYNLRVYCVDNTPCAFQEVVSDERHQGGQLLEEPKLLHFKGNTFSLQISV 470

Qy 688 HDVPSSLWKSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSD LACKLWVWQVEGD 747  
 |::| ||: | : ||::| :| :: ||| |::| |::| |::| : |::|  
 Db 471 LDIPPFLWRIKPFTACQEVFPFSRVWC SNRQPLHCAFSLERYTP TTTQLSCKICIRQLKGH 530

Qy 748 GQSF SINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGAD 807  
 | : :| : | : :: || || ||||: ||::| :: | | :| |  
 Db 531 EQILQVQTSILESERETITFFAQEDSTFPAQTGPKAFKIPYSIRQRICATFDT PNAKGKD 590

Qy 808 WRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHF PNGNLSQLAAAVAGLGQPDAGLFT 867  
 |: |||| :: :||:|:|: ||:|:|:|:|:|:| :|:| || |: :|: |  
 Db 591 WQMLAQKNSINRNLSYFATQSSPSAVILNLWEARHQHDGDLDSLACALEEIGRTHTKLSN 650

Qy 868 VSEAE 872  
 :||::  
 Db 651 ISESQ 655

RESULT 15

US-09-969-532-26

; Sequence 26, Application US/09969532  
 ; Patent No. 6777232  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Walke, D. Wade  
 ; APPLICANT: Scoville, John  
 ; TITLE OF INVENTION: No. 6777232el Human Membrane Proteins and Polynucleotides  
 Encoding the Same  
 ; FILE REFERENCE: LEX-0244-USA  
 ; CURRENT APPLICATION NUMBER: US/09/969,532  
 ; CURRENT FILING DATE: 2001-10-02  
 ; PRIOR APPLICATION NUMBER: US 60/237,280  
 ; PRIOR FILING DATE: 2000-10-02  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: FastSEQ for Windows Version 4.0  
 ; SEQ ID NO 26  
 ; LENGTH: 680  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 US-09-969-532-26

Query Match 26.5%; Score 1234; DB 4; Length 680;  
 Best Local Similarity 35.7%; Pred. No. 7.7e-108;  
 Matches 263; Conservative 128; Mismatches 215; Indels 130; Gaps 13;

Qy 197 VAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCASCGRGWQKRSRSC TNPAPLNGGAFC 256  
 :| ||||:| || |::|:| | |::| |||



Db 1 MAANIVAKRRSLSATVVVYVDGSWEVWSEWSVCSP----- 35

Qy 257 EGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DPAPRNGGEECQGTDLDT 316  
: | | | | | : | | | | : | : : :

Db 36 -----ECEHLRIRECTAPPPRNGGKFCEGLSQES 64

Qy 317 RNCTSDLCV-----HSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCRKKEGL 365  
| | | | : | : | | | | | : | : : : | | :

Db 65 ENCTDGLCILDKKPLHEIKPQSIENASDIALYSGL-GAAVVAVAVLVIGVTLYRRSQSDY 123

Qy 366 DSDVADSSILTSGFQPVSIKPSKA-----DNPHLL--TIQPDLTSTTTTTYQ 409  
| | | | | | : | : | | : | | | | : | |

Db 124 GVDVIDSSALTGGFQTFNFKTVRQAKNIMELMIQEKSFGNSLLNSAMQPDL-TVSRYS 182

Qy 410 GSLCPRQDGSPKFLTNHLLSPLG-----G 436  
| : | | | | : | : | | |

Db 183 GPIC-LQD-PLDKELMTESSLFNPLSDIKVKVQSSFMVSLGVSEAEYHGKNHSRTFPHG 240

Qy 437 GRHTLHHSSPTSEAEFFVSRLSTQNYFRSLPRGTSNMTYGTFFNLGGRIMIPNTGISILLI 496  
| : : | : : : | | | | | | | | : | | | | |

Db 241 NNHSFSTMHPRNKM-PYIQNLS-----SLPTRTELRTGVFGHLGGRLVMPNTGVSLLI 293

Qy 497 PPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE 556  
| | | : | | : : : | | | : | | | : | | : | | :

Db 294 PHGAIPEENSWEIYMSINQGEPLQSDGSEVLLSPEVTCGPPDMI VTPFALTIPHCAD 352

Qy 557 PSPDWSLRLKKQSCGSEWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 616  
| : : | | : : : | | : : : | | : | | : | : | | :

Db 353 VSSEHWNHLKRTQQGKWEEVMSVEDESTS--CYCLDPPFACHVLLDSFGTYALTGEPI 410

Qy 617 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 676  
: | | : : | : : | | : : | | : | | : | | : | | :

Db 411 TDCAVKQLKVAVFGCMSCNSLDYNLRVYCVDNTPCAFQEVVSDERHQGGQLLEEPKLLHF 470

Qy 677 KDSYHNLRLSIHDVPSSLWKSCLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 736  
| : : | : : | : | : | : | : | : | | : | | : | : | :

Db 471 KGNTFSLQISVLDIPPFLWRIKPFTACQEVFPFSRVWCSNRQPLHCAFSLERYTPTTTQLS 530

Qy 737 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS 796  
| | : : | : | | : : | : : | : : | | | | | : | | : | :

Db 531 CKICIRQLKGHEQILQVQTSILESERETITFFAQEDSTFPAQTGPKAFKIPYSIRQRICA 590

Qy 797 SLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVA 856  
: | | : | | : | | : : : | | : : | | : | | : | | :

Db 591 TFDTPNAKGKDWQMLAQKNSINRNLSYFATQSSPSAVILNLWEARHQHGDGDLDSLACALE 650

Qy 857 GLGQPDAGLFTVSEAE 872  
: | : | : | :

Db 651 EIGRTHTKLSNISESQ 666

Search completed: March 1, 2005, 09:05:50  
Job time : 56.3087 secs

OM protein - protein search, using sw model

Run on: March 1, 2005, 08:46:18 ; Search time 34.0799 Seconds  
(without alignments)  
2464.715 Million cell updates/sec

Title: US-10-624-932-2\_COPY\_26\_898  
Perfect score: 4660  
Sequence: 1 QQSATVANPVPGANPDLLPH.....AVAGLGQPDAGLFTVSEAEC 873

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

8						
Result	Query					
No.	Score	Match	Length	DB	ID	Description
1	977	21.0	919	2	T32541	unc-5 protein - Ca
2	977	21.0	947	1	B44294	unc-5 protein, lon
3	298.5	6.4	1584	2	T00026	brain-specific ang
4	296.5	6.4	1172	1	TSHUP2	thrombospondin 2 p
5	293	6.3	1074	2	JC5928	semaphorin F precu
6	293	6.3	1172	2	A42587	thrombospondin 2 p
7	276	5.9	1444	2	T18856	angiogenesis inhib
8	275	5.9	984	2	T00326	hypothetical prote
9	275	5.9	1522	2	T00028	brain-specific ang
10	274.5	5.9	1572	2	T00027	brain-specific ang
11	270.5	5.8	1170	2	A40558	thrombospondin 1 p
12	268.5	5.8	1170	1	TSHUP1	thrombospondin 1 p
13	263	5.6	1178	1	A39804	thrombospondin pre

14	243	5.2	469	1	S29126	properdin precurs
15	229	4.9	437	2	S05478	properdin - mouse
16	226	4.8	254	2	T15952	hypothetical prote
17	215	4.6	788	2	T25061	hypothetical prote
18	215	4.6	1651	2	T14160	transmembrane rece
19	208	4.5	1612	2	T30805	dutt1 protein - mo
20	205	4.4	1265	1	A37967	neural cell adhesi
21	188.5	4.0	957	2	T15976	hypothetical prote
22	188	4.0	1344	2	T14316	rig-1 protein - mo
23	187.5	4.0	1863	2	S46217	protein-tyrosine-p
24	186	4.0	423	2	T29549	hypothetical prote
25	181.5	3.9	1273	2	T42405	sax-3 protein - Ca
26	181	3.9	1736	2	A47747	tight junction pro
27	178	3.8	1745	2	A46431	tight junction-ass
28	172	3.7	837	2	T00355	hypothetical prote
29	171.5	3.7	1907	2	S50893	protein-tyrosine-p
30	169.5	3.6	934	1	A34372	complement C6 prec
31	168.5	3.6	152	2	D89753	protein F11C7.2 [i
32	166	3.6	860	2	T16892	hypothetical prote
33	162.5	3.5	654	2	T29247	hypothetical prote
34	159.5	3.4	951	2	T00017	gene ADAMTS-1 prot
35	159	3.4	805	2	T34212	hypothetical prote
36	158.5	3.4	2165	2	T21371	hypothetical prote
37	157.5	3.4	1501	2	I58148	protein-tyrosine-p
38	157	3.4	550	2	T47158	hypothetical prote
39	156.5	3.4	807	2	A38152	F-spondin - rat
40	156	3.3	584	1	C8HUA	complement C8 alph
41	153.5	3.3	1898	2	S46216	leukocyte antigen-
42	150.5	3.2	1437	2	T31093	probable protein-t
43	149.5	3.2	712	2	A45638	immunodominant mic
44	148.5	3.2	206	2	A45517	coccidiosis-relate
45	148.5	3.2	590	2	I46687	complement compone

#### ALIGNMENTS

##### RESULT 1

T32541

unc-5 protein - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T32541

R;Latreille, P.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of C. elegans cosmid B0273.

A;Reference number: Z21187

A;Accession: T32541

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-919 <LAT>

A;Cross-references: UNIPROT:O44171; EMBL:AF036698; PIDN:AAB88355.1;

GSPDB:GN00022; CESP:B0273.4a

A;Experimental source: strain Bristol N2; clone B0273

C;Genetics:

A;Gene: unc-5; CESP:B0273.4a

A;Map position: 4

A;Introns: 41/3; 108/1; 142/3; 201/1; 323/2; 553/1; 858/3  
 C;Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology;  
 thrombospondin type 1 repeat homology

Query Match 21.0%; Score 977; DB 2; Length 919;  
 Best Local Similarity 28.7%; Pred. No. 1.1e-62;  
 Matches 265; Conservative 168; Mismatches 379; Indels 110; Gaps 31;

Qy	24	EPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIER--STDGSSGLPTMEVRI	81
		: :  ::   :    ::    :  :::  :     :    :  : :: :	
Db	9	QPKSGYVIRNKPLRLQCRANHATKIRYKCSSKWID--DSRIEKLIGTDSTSGVGIDASV	66
Qy	82	NVSRQQVEKVFGLLEEYWCQCVAWSSSG-----TTKSQKAYIRIARLRKNFEQEPLAKEVS	136
		::    : :: :            :  :  ::  :  : :	
Db	67	DISRIDVDTSGHVDAFQCQCYA---SGDDQDVVASDVATVHLAYMRKHFLKSPVAQRVQ	123
Qy	137	LEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTC	196
		:    :        :   :: :         :   ::    : :	
Db	124	EGTTLQLPCQAPESDPKAELTWYKDGVVVQP--DANVIRASDGLSIMSAARLSDSGNYTC	181
Qy	197	VAKNIVARRRSASAAVIVYVNGGWSTWTEW-SVCSASCG-----RGWQKR	240
		:  : :  ::     :	
Db	182	EATNVANSRKTDPEVQIYVDGGWSESWPWIGTCHVDCPLLRQHAHRIRDPHDVLPHQRR	241
Qy	241	SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWRSDP	300
		: :          : ::: :    :              : :   :	
Db	242	TRTCNNPAPLNDGEYCKGEEEMTRSKVPCKLDGGWSSWSWSACSSSCHRYRTRACTVP	301
Qy	301	APRNGGEECQGTDLDRNCTSDLCVHSASG--PEDVALYVGLIAVAVCLVLLLLVLILVY	358
		:         :   :    :     :  : :: : :  :  :	
Db	302	PPMNGGQPCFGDDLMTQECPAQLCTADSSRIVISDTAVYGSVASIFIVASFILAILAMFC	361
Qy	359	CR-----KKEGLSDSDVADSSILTSQFQVPSIKPSKADNPHELLTI-----	397
		:   : :: :   : :  : : : :  :	
Db	362	CKRGNSKSKPLKPQKMNSEKAGGIYYS---EPPGVRRLLEHQHGTLLGEKISSCSQYF	418
Qy	398	-QPDLSSTTT-----TYQGSCLPRQDGPSPKFQLTNGHLLSPLGGGRHTLHSSPT-SE	449
		:   :	
Db	419	EPPPLPHSTTLRSGKSAFSGYSSTRNAGSRAALIQECSSSSSSGSGGKRTMLRTSSSNCSD	478
Qy	450	AEEFVSRLSTQNYFRSLPRGTS-NMTYGTENFLGGRMLIPNTGISLLIPPDAIPRGKIYE	508
		: : :       : :      : :   ::   :  :	
Db	479	DDNYATLYDYMEDKSVLGLDTSQNIVAAQIDSNGARLSLSKSGARLIVPELAVEGEKM--	536
Qy	509	IYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGV-----LLTRPVILAMDHCGEPSP-D	560
		:   : :   ::    :     :     :::	
Db	537	LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVSMASHDNILRRPVVVSFRHCASTFPRD	596
Qy	561	SWSRLKKQSCEGS-WEDVLHLGEEAPSHLYYCQLEASA-----CYVFTEQLGRFAL	611
		:    :      : : :    : :	
Db	597	NWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQFEQPGKKNDGFGWCHVMTYSLARLML	654
Qy	612	VGEAL--SVAAAKRLKLLLFPVACTSLE--YNIRVYCLHDTHDALKEVVQLEKQLGGQL	667
		::   :   :    : : :   :    ::   :   :	
Db	655	AGHPRRNSLSAAKRVLAVFGPTMSAYRRPFELRVYCVPETGAAMESVWKQED--GSRL	712
Qy	668	IQEPR--VLHFKDSYHNLRSLSIHDV-PSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFT	724

```

      : |      :|: |      || : | || |      :| |      :      | |||:
Db      713 LCESNDFILNEKG---NLCICIEDVIPGFSCDGPVEVVEISETQHRFV---AQNGLHCSLK 766

Qy      725 LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAF 784
      : |      : :: | :|      :      : :: :      : : | | |      :      |
Db      767 FRPKEINGSQFSTRVIVYQKASSTEPMM--EVSNEPELYDATSEEREKGSVCV----EF 820

Qy      785 KIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKP--SPTAMILNLWEARH 842
      ::|| :: ::      || |      :||| ||:|||| | :| |||| |      |||:::|:||||
Db      821 RLPFGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFPDCSPTSLLLLDLWEASS 880

Qy      843 FPNGN-LSQLAAAVAGLGQPDa 863
      :      : |      : :|:||||
Db      881 SGSARAVPDLLQTLRVMGRPDA 902

```

## RESULT 2

B44294

unc-5 protein, long form - *Caenorhabditis elegans*

N;Contains: unc-5 protein, short form

C;Species: *Caenorhabditis elegans*

C;Date: 30-Apr-1993 #sequence\_revision 28-Jul-1995 #text\_change 09-Jul-2004

C;Accession: B44294; T32540; A44294

R;Leung-Hagesteijn, C.; Spence, A.M.; Stern, B.D.; Zhou, Y.; Su, M.W.; Hedgecock, E.M.; Culotti, J.G.

Cell 71, 289-299, 1992

A;Title: UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 domains, guides cell and pioneer axon migrations in *C. elegans*.

A;Reference number: A44294; MUID:93046629; PMID:1384987

A;Contents: variety Bergerac

A;Accession: B44294

A;Molecule type: DNA

A;Residues: 1-947 <LEU>

A;Cross-references: UNIPROT:O44171; GB:S47168; NID:g258527; PIDN:AAB23867.1; PID:g258529

A;Note: sequence extracted from NCBI backbone (NCBIN:116668, NCBIN:116670, NCBIN:116672, NCBIN:116674, NCBIN:116676, NCBIN:116678, NCBIN:116680, NCBIN:116682, NCBIN:116685, NCBIP:118648)

A;Note: authors translated the codon CTA for residue 642 as Val; sequence shown follows the authors' translation.

A;Note: mRNA lacking the first exon is equally prevalent

R;Latreille, P.

submitted to the EMBL Data Library, December 1997

A;Description: The sequence of *C. elegans* cosmid B0273.

A;Reference number: Z21187

A;Accession: T32540

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-947 <LAT>

A;Cross-references: EMBL:AF036698; PIDN:AAB88356.1; GSPDB:GN00022; CESP:B0273.4b

A;Experimental source: strain Bristol N2; clone B0273

C;Genetics:

A;Gene: unc-5

A;Map position: 4

A;Introns: 28/1; 69/3; 136/1; 170/3; 229/1; 351/2; 581/1; 886/3

C;Function:



Qy 509 IYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGV-----LLTRPVILAMDHCGEPS-D 560  
 :|| : : | : : ||| : : | | : : ||| : : | |  
 Db 565 LYLAVSDTLTDQPHLKPIESALSPVIVIGQCDVSMASHDNILRRPVVVSFRHCASTFFRD 624

Qy 561 SWSLRLKKQSCEGS-WEDVLHLGEEAPSHLYYCQLEASA-----CYVFTEQLGRFAL 611  
 :| | : ||| : : :||| : : | | | :| | | | |  
 Db 625 NWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQFEQPGKKNDGFGWCHVMTYSLARLML 682

Qy 612 VGEAL--SVAAAKRLKLLLFAPVACTSLE--YNIRVYCLHDTHDALKEVVQLEKQLGGQL 667  
 | :||| : | :| | : : :||| : :| | :| : | :|  
 Db 683 AGHPRRNSLSAAKRVHLAVFGPTMSAYRRPFELRVYCVPETGAAMESVWKQED--GSRL 740

Qy 668 IQEPR--VLHFKDSYHNRLRLSIHDV-PSSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFT 724  
 : | :| : | | : | | | : | | : | | | :  
 Db 741 LCESNDFILNEKG---NLCICIEDVIPGFSCDGPVEVEISETQHRETV---AQNGLHC SLK 794

Qy 725 LERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTREAEALLALESEAGVPALVGPSAF 784  
 : | : : :| : : : : : : : | | | : |  
 Db 795 FRPKEINGSQFSTRVIVYQKASSTEPMMV--EVSNEPELYDATSEEREKGSVCV----EF 848

Qy 785 KIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKP--SPTAMILNLWEARH 842  
 :|| : : : || | :||| ||:||| | :| |||| | ||| : :| :|||  
 Db 849 RLPFGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFPDCSPTSLLLDLWEASS 908

Qy 843 FPNGN-LSQLAAAVAGLGQPD 863  
 : : | : :| :|||  
 Db 909 SGSARAVPDLLQTLRVMGRPDA 930

# RESULT 3

T00026

brain-specific angiogenesis inhibitor 1 - human

N;Alternate names: BAI1 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 09-Jul-2004

C;Accession: T00026

R;Nishimori, H.; Shiratsuchi, T.; Urano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshida, S.; Ono, M.; Kuwano, M.; Nakamura, Y.

submitted to the EMBL Data Library, June 1997

A;Reference number: Z14064

A;Accession: T00026

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1584 <NIS>

A;Cross-references: UNIPROT:O14514; EMBL:AB005297; NID:d1175078; PID:d1024528

A;Experimental source: brain

C;Genetics:

A;Gene: GDB:BAI1

A;Cross-references: GDB:9838088; OMIM:602682

A;Map position: 8q24-8q24

F;408-462/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 6.4%; Score 298.5; DB 2; Length 1584;

Best Local Similarity 33.5%; Pred. No. 4.1e-13;

Matches 78; Conservative 35; Mismatches 91; Indels 29; Gaps 11;

Qy 99 CQCVAWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVIEW 158

```

      | | :| | : :| | : | | | | |
Db      309 CNREACGPAGRTSSRSQSLRSTDARR---REELGDEL---QQFGFPA-PQTGDPAAE-EW 360
Qy      159 LRNEDLVDP SLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 218
      : | | : || | : ::| : : : | :|
Db      361 --SPWSVCSSTCGEGWQTR-----TRFCVSSSYSTQCSGPLREQRLCNSAVCPVHG 410
Qy      219 GWSTWTEWSVCSASCGRGWQKRSRSCNPNAPLNGGAFCEGQNVQKTAC-ATLCP---VDG 274
      | | : || : || : || : | : | | | | | | | | | |
Db      411 AWDEWSPWSLCSSSTCGRGFRDRTRTCR--PPQFGGNPCEGPEKQTKFCNIALCPGRAVDG 468
Qy      275 SWSPWSKWSACGLDCT---HWSRRECSDPAPRNGGEECQGTDLDTNRCTSDLC 324
      :| : || | | | | : | : | : | : | | | | :| :| |
Db      469 NWNEWSSWSACSASCSQGRQQRTRRECNGPS--YGGAECQGHVETRD CFLQQC 519

```

#### RESULT 4

##### TSHUP2

thrombospondin 2 precursor - human

C;Species: Homo sapiens (man)

C;Date: 19-May-1995 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004

C;Accession: A47379; A42173

R;LaBell, T.L.; Byers, P.H.

Genomics 17, 225-229, 1993

A;Title: Sequence and characterization of the complete human thrombospondin 2

cDNA: potential regulatory role for the 3' untranslated region.

A;Reference number: A47379; MUID:94010892; PMID:8406456

A;Accession: A47379

A;Molecule type: mRNA

A;Residues: 1-1172 <LAB>

A;Cross-references: UNIPROT:P35442; GB:L12350; NID:g307505; PIDN:AAA03703.1;

PID:g307506

R;LaBell, T.L.; Milewicz, D.J.; Disteché, C.M.; Byers, P.H.

Genomics 12, 421-429, 1992

A;Title: Thrombospondin II: partial cDNA sequence, chromosome location, and

expression of a second member of the thrombospondin gene family in humans.

A;Reference number: A42173; MUID:92217961; PMID:1559694

A;Accession: A42173

A;Molecule type: mRNA

A;Residues: 560-1172 <LA2>

A;Cross-references: GB:M81339

A;Experimental source: fibroblast

A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIP:95096)

C;Genetics:

A;Gene: GDB:THBS2; TSP2

A;Cross-references: GDB:128789; OMIM:188061

A;Map position: 6q27-6q27

C;Complex: homotrimer, disulfide linked

C;Function:

A;Description: participates in cell migration and adhesion, and in platelet aggregation

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer

F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-1172/Product: thrombospondin 2 #status predicted <MAT>



F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>  
 F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>  
 F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>  
 F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>  
 F;553-588/Domain: EGF homology <EGF1>  
 F;652-691/Domain: EGF homology <EGF>  
 F;928-930/Region: cell attachment (R-G-D) motif  
 F;151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent)  
 #status predicted  
 F;167-226/Disulfide bonds: #status predicted  
 F;266,270/Disulfide bonds: interchain #status predicted  
 F;612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 6.4%; Score 296.5; DB 1; Length 1172;  
 Best Local Similarity 30.5%; Pred. No. 3.8e-13;  
 Matches 78; Conservative 28; Mismatches 105; Indels 45; Gaps 9;

```

Qy      184 RQARLADTANYTCVAKNIVARRRSASAA-VIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 242
      :: | | : ||: :| | | | : :||| |: || || :|| | | |
Db      403 QRGRSCDVTSNTCLGPSIQTRACSLSKCDTRIRQDGGWSHWSPWSSCSVTCGVGNITRIR 462

Qy      243 SCTNPAPLNGGAFCEGQNVOQTAC-ATLCPVDGWSWSPWSKWSACGLDCT---HWRSRECS 298
      | :| | || |:| : || ||:| | || | || | : | | :| | :
Db      463 LCNSPVPQMGGKNCKGSGRETKACQGAPCPIDGRWSPWSPWSACTVTCAGGIRERTRVCN 522

Qy      299 DPAPRNGGEECQGTDLDRNCTSDLCVHSASGPDVALYVGLIAVAVCLVLLLLVLILVY 358
      | | : ||: | | : : | | | | | | | | | |
Db      523 SPEPQYGGKACVGDVQERQMCNKRSC-----PVDGCLSNPCFPGAQC----- 564

Qy      359 CRKKEGLDSDVADSSILTSGFQPVSI--KPSKADNPHELLTIQPDLSSTTTT-----TYQ 409
      | | | : || || : : : : : ||: :|: |
Db      565 -----SSFPDGS-WSCGFPCVGFILNGTHCEDLDECALVPDICFSTSKVPRCVNTQP 615

Qy      410 GSLC----PRQDGPSP 421
      | | || | |
Db      616 GFHCLPCPPRYRGNQP 631
  
```

# RESULT 5

JC5928

semaphorin F precursor - human

C;Species: Homo sapiens (man)

C;Date: 10-Apr-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004

C;Accession: JC5928

R;Simmons, A.D.; Pueschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.  
 Biochem. Biophys. Res. Commun. 242, 685-691, 1998

A;Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-  
 chat candidate interval.

A;Reference number: JC5928; MUID:98125554; PMID:9464278

A;Accession: JC5928

A;Status: nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-1074 <SIM>

A;Cross-references: UNIPROT:Q13591; GB:U52840; NID:g2772583; PIDN:AAC09473.1;  
 PID:g2772584

A;Experimental source: brain

C;Comment: This protein disrupts normal brain development and leads to some of the features of Cri-du-chat.

C;Genetics:

A;Gene: sema4

C;Superfamily: human semaphorin F; thrombospondin type 1 repeat homology

F;1-20/Domain: signal sequence #status predicted <SIG>

F;50-533/Domain: semaphorin #status predicted <SEM>

F;840-896/Domain: thrombospondin type 1 repeat homology <THR3>

F;971-993/Domain: transmembrane #status predicted <TMM>

Query Match 6.3%; Score 293; DB 2; Length 1074;

Best Local Similarity 45.8%; Pred. No. 6.1e-13;.

Matches 54; Conservative 11; Mismatches 49; Indels 4; Gaps 2;

```
Qy      216 VNGGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACATL-CPVDG 274
          ||| || || || || | || : | | | | | || | | | | | | | | | | |
Db      783 VNGAWSAWTSWSQCSRDCSRGIRNRKRVCNNPEPKYGGMPCLGPSLEYQECNTLPCPVDG 842

Qy      275 SWSPWSKWSACGLDC---THWRSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSAS 329
          || || |: | | : | |: | | | | | | | | | | | | | | | | |
Db      843 VWSCWSPWTKCSATCGGGHYMRTRSCSNPAPAYGGDICLGLHTEELCNTQPCPESWS 900
```

#### RESULT 6

A42587

thrombospondin 2 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C;Accession: A42587; A39851

R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.  
J. Biol. Chem. 267, 3274-3281, 1992

A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cell growth and development.

A;Reference number: A42587; MUID:92147683; PMID:1371115

A;Accession: A42587

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: nucleic acid

A;Residues: 1-1172 <LAH>

A;Cross-references: UNIPROT:Q03350; GB:L07803; GB:M87275; NID:g340421;  
PIDN:AAA53064.1; PID:g567241

A;Note: sequence extracted from NCBI backbone (NCBIP:81502)

R;Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.;  
Dixit, V.M.

J. Biol. Chem. 266, 12821-12824, 1991

A;Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.

A;Reference number: A39851; MUID:91302287; PMID:1712771

A;Accession: A39851

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-873 <BOR>

A;Cross-references: GB:M64866; NID:g201994; PIDN:AAA40432.1; PID:g201995

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

C;Keywords: calcium binding; glycoprotein

F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>

F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>

F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>  
 F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>  
 F;553-588/Domain: EGF homology <EGF1>  
 F;652-691/Domain: EGF homology <EGF>

Query Match 6.3%; Score 293; DB 2; Length 1172;  
 Best Local Similarity 38.0%; Pred. No. 6.9e-13;  
 Matches 60; Conservative 22; Mismatches 66; Indels 10; Gaps 5;

```

Qy      184 RQARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSR 242
          :: | | : ||: :| | | : |||| | : || || :|| | | |
Db      403 QRGRSCDVTSNTCLGPSIQTRTCSLGTKCDTRIRQNGGWSHWPWSSCSVTGCVGNVTRIR 462

Qy      243 SCTNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGWSWSPWSKWSACGLDCT---HWRSRECS 298
          | :| | || | :| : | ||:|| |||| |||| : | || | :
Db      463 LCNSPVPQMGGKNCKGSGRETKPCQRDPCPIDGRWSPWSPWSACTVTCAGGIRERSRVCN 522

Qy      299 DPAPRNGGEECQG--TD---LDTRNCTSDLCVHSASGP 331
          | | : ||:| | | : : | : | | : : |
Db      523 SPEPQYGGKDCVGDVTEHQMCNKRSCPIDGCLSNPCFP 560
  
```

# RESULT 7

T18856

angiogenesis inhibitor homolog - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T18856; T24653

R;McMurray, A.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z19031

A;Accession: T18856

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1444 <WIL>

A;Cross-references: UNIPROT:Q8MYA8; EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1

A;Experimental source: clone C02B4

R;McMurray, A.

submitted to the EMBL Data Library, July 1995

A;Reference number: Z19917

A;Accession: T24653

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1444 <WI2>

A;Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1

A;Experimental source: clone T07C5

C;Genetics:

A;Gene: CESP:C02B4.1

A;Map position: X

A;Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566/2; 625/1; 696/2; 786/3; 812/2; 878/3; 971/1; 1007/3; 1067/1; 1099/3; 1180/3; 1273/2; 1305/1; 1363/1; 1388/2

Query Match 5.9%; Score 276; DB 2; Length 1444;  
 Best Local Similarity 27.2%; Pred. No. 1.6e-11;  
 Matches 73; Conservative 28; Mismatches 97; Indels 70; Gaps 12;

```

Qy      98 WCQCVAWSSSGTTKSQKAYIRIARLRKNFEQ-----EPLAKEVSLEQGIVLPCRPPEGI 151
      | : :||:      : : | : | |      :      ||| :|| |
Db      1134 WSEWSSWSAC-----SCFSLTSTRRRRFCQVVDPTVQGFCAGAILAQ---IPCAPGSCS 1183

Qy      152 PPAE-----VEW-----LRNEDLVDPNVDPNVYITREHSLVVRQARLADTAN 193
      | |      ||      :|| :|      :      | | :
Db      1184 PSAGGWSLWSEWSSCSKDCGDTGHQIRNRCSEP-----IPSNRGAYCSG 1228

Qy      194 YT-----CVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNPA 248
      |:      || |: : :      |:|||: || || |: | | : |:| | ||
Db      1229 YSFDQRPCVMDNVCSDEK-----VDGGWTDWTAWSECTDYCRNGHRSRTRFCANPK 1279

Qy      249 PLNGGAFCEGQNVQKTAC--ATLCPV-DGSWSPWSKWSACGLDC---THWRSRECSDPAP 302
      | ||| | | : : |      | : || || || |: | |      | ||| | |
Db      1280 PSQGAQCTGSDFELNPCFDPARCHLRDGGWSTWSDWTPCSASCGFGVQTRDRSCSSPEP 1339

Qy      303 RNGGEECQGTDLDRNCTSDLCVHSASG 330
      : ||: | |      | |      | | : |
Db      1340 K-GGQSCSGLAHQTSLCDLPACDHESDG 1366

```

RESULT 8

T00326

hypothetical protein KIAA0550 - human

C;Species: Homo sapiens (man)

C;Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 15-Mar-2004

C;Accession: T00326

R;Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.

DNA Res. 5, 31-39, 1998

A;Title: Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.

A;Reference number: Z14086; MUID:98290545; PMID:9628581

A;Accession: T00326

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-984 <NAG>

A;Cross-references: EMBL:AB011122; NID:g3043623; PIDN:BAA25476.1; PID:g3043624

A;Experimental source: brain

C;Genetics:

A;Note: KIAA0550

F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 5.9%; Score 275; DB 2; Length 984;

Best Local Similarity 39.0%; Pred. No. 1.1e-11;

Matches 57; Conservative 20; Mismatches 53; Indels 16; Gaps 6;

```

Qy      195 TCVA-----KNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNPA 248
      |||:      : | : : |:| | |: ||:| | :||| : |:| |||
Db      317 TCVSPYGTHCSGFLRESRVCNNALCPVHGVWEEWSPWSLCSFTCGRGQRTTRTRSCT--P 374

Qy      249 PLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDC---THWRSRECSDPAPRN 304
      | || |||      | ||||| | || || | : | |      |||:|: | :
Db      375 PQYGGRPCEGPETHHKPCNIALCPVDGQWQEWSSWSQCSVTCSNGTQQRSRQCT--AAAH 432

```

Qy 305 GGEQCQGTDLDTNRCTSDLCVHSASG 330  
 || ||:| :| | : | :|:|  
 Db 433 GGSECRGPWAESRECYNPEC--TANG 456

RESULT 9

T00028

brain-specific angiogenesis inhibitor 3 - human

N;Alternate names: BAI3 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 09-Jul-2004

C;Accession: T00028

R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.

Cytogenet. Cell Genet. 79, 103-108, 1997

A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI 1).

A;Reference number: Z14066; MUID:98194217; PMID:9533023

A;Accession: T00028

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-1522 <SHI>

A;Cross-references: UNIPROT:O60242; EMBL:AB005299; NID:g3021700;

PIDN:BAA25363.1; PID:g3021701

A;Experimental source: brain

C;Genetics:

A;Gene: GDB:BAI3

A;Cross-references: GDB:9838090; OMIM:602684

A;Map position: 6q12-6q12

F;344-398/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 5.9%; Score 275; DB 2; Length 1522;  
 Best Local Similarity 39.0%; Pred. No. 2e-11;  
 Matches 57; Conservative 20; Mismatches 53; Indels 16; Gaps 6;

Qy 195 TCVA-----KNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNPA 248  
 |||: : | : : | : ||| : |||| : | : ||||  
 Db 317 TCVSPYGTHCSGPLRESRVCNNTALCPVHGVWEEWSPWSLCSFTCGRGQRTTRRSCT--P 374  
 Qy 249 PLNGGAFCEGQNQKTAC-ATLCPVDGWSWPWSKWSACGLDC---THWRSRECSDPAPRN 304  
 | || ||| | ||||| | || || : | | |||:| : | :  
 Db 375 PQYGGRPCEGPETHHKPCNIALCPVDGQWQEWSSWSQCSVTCSNGTQQRSRQCT--AAAH 432  
 Qy 305 GGEQCQGTDLDTNRCTSDLCVHSASG 330  
 || ||:| :| | : | :|:|  
 Db 433 GGSECRGPWAESRECYNPEC--TANG 456

RESULT 10

T00027

brain-specific angiogenesis inhibitor 2 - human

N;Alternate names: BAI2 protein

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 09-Jul-2004

C;Accession: T00027

R;Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.

Cytogenet. Cell Genet. 79, 103-108, 1997

A;Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain-specific angiogenesis inhibitor 1 (BAI 1).  
 A;Reference number: Z14066; MUID:98194217; PMID:9533023  
 A;Accession: T00027  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-1572 <SHI>  
 A;Cross-references: UNIPROT:O60241; EMBL:AB005298; NID:g3021698;  
 PIDN:BAA25362.1; PID:g3021699  
 A;Experimental source: brain  
 C;Genetics:  
 A;Gene: GDB:BAI2  
 A;Cross-references: GDB:9838089; OMIM:602683  
 A;Map position: lp35-lp35

Query Match 5.9%; Score 274.5; DB 2; Length 1572;  
 Best Local Similarity 19.2%; Pred. No. 2.2e-11;  
 Matches 176; Conservative 108; Mismatches 307; Indels 327; Gaps 38;

Qy	148	PEGIPPAEVEWLRNEDLVDPSPDNVY-----ITREHSLVVRQARL	188
		: :    :   :  :  :	
Db	271	PEEPPKVKQTQWPRSAD-----EPGLYMAQTGDPAAEEWSPWSVCSLTCGQGLQVR-TRS	323
Qy	189	ADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSTNPA	248
		::   : :   : :   :        :        :     :	
Db	324	CVSSPYGTLCSGPLRETRPCNNSATCPVHGVWEEWGSWSLCSRSCGRGSRSRMRTCV--P	381
Qy	249	PLNGGAFCEGQNVQKTACA-TLCPVDGWSWPWSKWSACGLDC---THWRSRECSDPAPR-	303
		:       :    :    :                   :	
Db	382	PQHGGKACEGPELQTKLCSMAACPVEGQWLEWGPWGPCSTSCANGTQQRSRKCSVAGPAW	441
Qy	304	-----NGGEECQ	310
		:	
Db	442	ATCTGALTDTRECSNLECPATDSKWGPWNAWSLCSKTCDTGWQRRFRMCQATGTQGYPCE	501
Qy	311	GTDLDRNCTSDLC--VHSASGPEDVAL-----	336
		: :   :	
Db	502	GTGEEVKPCSEKRCPAFHEMCRDEYVMLMTWKKAAAAGEIIYNKCPPNASGSASRRCLLSA	561
Qy	337	----YVGLIAVAVCL---VLLLLVLILVYCRKKEGLDSDVADSSILTSGFQPVSIKPSKA	389
		:     :   : : :   : : :   : :     : : :	
Db	562	QGVAYWGLPSFARCISHEYRYLYLSLREHLAKGQRLAGEGMSQVVRS-LQELLARRTYY	620
Qy	390	DNPHLLTIQPDLSSTTTTTYQGSCLPRQDGPSPKFQLT-----NGHLLSPLGG	436
		: : :     : :        : : :   :	
Db	621	SGDLLFSVDILRNVTDTFKRATYVPSADDVQRFQVVSEFMVDAENKEKWDDAQQVSP--G	678
Qy	437	GRHTLHHSSPTSEAEFEV-----SRLSTQNYFRSLPRG-----TSNMTYGTFN	479
		: :   :       :   : : :	
Db	679	SVHLLR-----VVEDFIHLVGDAKAFQSSLIVTDNLVISIQREPVSASVSSDITFPMRG	732
Qy	480	FLG-----GRLMIPNTGISLLIP-----PDAIPRGK-----	505
		:   :          :	
Db	733	RRGMKDWVRHSEDRFLPKVELSLSSPGKPATSGAAGSPGRGRGPGTVPPGPGHSHQRL	792
Qy	506	-----IYE-IYLTLHKPEDVRLPLAGCQTLTLLSPIVSCGPPGVLLTRPVIL	549
		:   :             : : :   :      :	

Db 793 PADPDESSYFVIGAVLYRTLGLILPPP---RPPLAVTSRVMT--VTVRPPTQPPAEPLIT 847  
 Qy 550 A-----MDHCGEPPSPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQ-LEASACYV-- 601  
       :: :| || : : | | : || || | :  
 Db 848 VELSYIINGTTDPHCASWDYS-RADASSGDWD-----TENCQTLETQAAHTRC 894  
 Qy 602 FTEQLGRFALVGE-----ALSVAALKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEV 656  
       : | ||:: : | : : : |:: | : | : : | |  
 Db 895 QCQHLSTFAVLAQPPKDLTLELAGSPSVPLVIGCAVSCMALLTLAIYA-----AFWRF 948  
 Qy 657 VQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDPSSLWKSLLVSYQEIPFYHIWNGTQ 716  
       :: | : : | ||| : |:: : || : : |  
 Db 949 IKSERSI-----ILLNFCLSI--LASNI---LILVGQSRVLSKGVCTMTA 988  
 Qy 717 RYLHCTFTLERVSPSTSDLACKLWV-----WQVEGDG 748  
       :|| | | : || | | :  
 Db 989 AFLHFFF-----LSSFCWVLTEAWQSYLAVIGRMTRLVRKRFLCLGWGLPALV 1037  
 Qy 749 QSFSINFNITKDTRFAELLALESEAG-VPALVGPSA-----FKIPFLIRQKI-----IS 796  
       : | : | || : | | | : | ||| : | : : | : ||  
 Db 1038 VAVSVGFTRTKGYGTSSYCWLSLEGGLLYAFVGPAAVIVLVNMLIGIIVFNKLMARDGIS 1097  
 Qy 797 SLDPPCRRGAD---WRTL 811  
       | |:: | : |  
 Db 1098 DKSKKQRAGSERCPWASL 1115

# RESULT 11

A40558

thrombospondin 1 precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence\_revision 05-Jun-1992 #text\_change 09-Jul-2004

C;Accession: A40558; A37905; B42587; S68787

R;Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.

Genomics 11, 587-600, 1991

A;Title: Characterization of the murine thrombospondin gene.

A;Reference number: A40558; MUID:92128941; PMID:1774063

A;Accession: A40558

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1170 <LAW>

A;Cross-references: UNIPROT:P35441; GB:M62449; GB:M62450; GB:M62451; GB:M62452;  
 GB:M62453; GB:M62454; GB:M62455; GB:M62456; GB:M62457; GB:M62458; GB:M62459;  
 GB:M62460; GB:M62461; GB:M62462; GB:M62463; GB:M62464; GB:M62465; GB:M62466;  
 GB:M62467; GB:M62468; GB:M62469; GB:M62470; NID:g511867; PIDN:AAA50611.1;  
 PID:g511869

R;Bornstein, P.; Alfi, D.; Devarayalu, S.; Framson, P.; Li, P.  
 J. Biol. Chem. 265, 16691-16698, 1990

A;Title: Characterization of the mouse thrombospondin gene and evaluation of the  
 role of the first intron in human gene expression.

A;Reference number: A37905; MUID:90375546; PMID:2398070

A;Accession: A37905

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-490 <BOR>

A;Cross-references: GB:J05605; GB:J05606; NID:g201991; PIDN:AAA40431.1;  
 PID:g554390  
 R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.  
 J. Biol. Chem. 267, 3274-3281, 1992  
 A;Title: Characterization of mouse thrombospondin 2 sequence and expression  
 during cell growth and development.  
 A;Reference number: A42587; MUID:92147683; PMID:1371115  
 A;Accession: B42587  
 A;Status: preliminary; not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-1152,'P',1154-1170 <LAH>  
 A;Cross-references: GB:M87276  
 A;Note: sequence extracted from NCBI backbone (NCBIP:81501)  
 R;Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.  
 FEBS Lett. 387, 36-41, 1996  
 A;Title: Expression and initial characterization of recombinant mouse  
 thrombospondin 1 and thrombospondin 3.  
 A;Reference number: S68787; MUID:96234006; PMID:8654563  
 A;Accession: S68787  
 A;Molecule type: protein  
 A;Residues: 19-26,'X',28-37 <CHE>  
 C;Complex: homotrimer, disulfide linked  
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat  
 homology; von Willebrand factor type C repeat homology  
 C;Keywords: calcium binding; glycoprotein; homotrimer  
 F;1-18/Domain: signal sequence #status predicted <SIG>  
 F;19-1170/Product: thrombospondin 1 #status predicted <MAT>  
 F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>  
 F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>  
 F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>  
 F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>  
 F;551-586/Domain: EGF homology <EGF>  
 F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.8%; Score 270.5; DB 2; Length 1170;  
 Best Local Similarity 32.2%; Pred. No. 3e-11;  
 Matches 57; Conservative 24; Mismatches 71; Indels 25; Gaps 5;

```

Qy      182 VVRQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGGWSTWTEWSVCSASC 233
          : :: | | : | | :: | : | : || | | : | | | : |
Db      399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWSPWSSCSVTC 451

Qy      234 GRGWQKRSRSCNTPAPLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDC--- 289
          | | | | | : | : | | | : : | | | : | | | | | : |
Db      452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPSPWDICSVTCGGG 511

Qy      290 THWRSRECSDPAPRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGLIAVAVC 346
          || | : : | | : | : | | : : | | | | | | | | |
Db      512 VQRRSRLCNNPTPQFGGKDCVGDVTENQVCNKQDC-----PIDGCLSNPCFAGAKC 562

```

# RESULT 12

TSHUP1

thrombospondin 1 precursor - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1987 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004

C;Accession: A26155; A34274; A30140; A25812; A05172; A42927



R;Lawler, J.; Hynes, R.O.  
 J. Cell Biol. 103, 1635-1648, 1986  
 A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.  
 A;Reference number: A26155; MUID:87057617; PMID:2430973  
 A;Accession: A26155  
 A;Molecule type: mRNA  
 A;Residues: 1-1170 <LAW>  
 A;Cross-references: UNIPROT:P07996; GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138  
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing  
 R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.  
 J. Biol. Chem. 264, 11222-11227, 1989  
 A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.  
 A;Reference number: A34274; MUID:89291870; PMID:2544587  
 A;Accession: A34274  
 A;Molecule type: DNA  
 A;Residues: 1-166 <LAH>  
 A;Cross-references: GB:J04835  
 R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.  
 J. Cell Biol. 108, 729-736, 1989  
 A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.  
 A;Reference number: A30140; MUID:89139590; PMID:2918029  
 A;Accession: A30140  
 A;Molecule type: mRNA  
 A;Residues: 1-83,'A',85-522,'A',524-1170 <HEN>  
 A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465  
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing  
 R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.  
 Biochemistry 25, 8418-8425, 1986  
 A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.  
 A;Reference number: A25812; MUID:87157592; PMID:3030396  
 A;Accession: A25812  
 A;Molecule type: mRNA  
 A;Residues: 1-83,'A',85-397 <KOB>  
 A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354  
 R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986  
 A;Reference number: A05172; MUID:86287276; PMID:3461443  
 A;Accession: A05172  
 A;Molecule type: mRNA  
 A;Residues: 1-83,'A',85-374,'RC' <DIX>  
 A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801  
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing  
 R;Sun, X.; Skorstengaard, K.; Mosher, D.F.  
 J. Cell Biol. 118, 693-701, 1992  
 A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.  
 A;Reference number: A42927; MUID:92348511; PMID:1379247  
 A;Accession: A42927

A;Molecule type: protein  
 A;Residues: 987-1003 <SUN>  
 A;Note: Cys-992 is shown to have a free sulfhydryl  
 C;Genetics:  
 A;Gene: GDB:THBS1; TSP1; TSP  
 A;Cross-references: GDB:120438; OMIM:188060  
 A;Map position: 15q15-15q15  
 A;Introns: 23/1  
 A;Note: the list of introns may be incomplete  
 C;Complex: homotrimer, disulfide linked  
 C;Function:  
 A;Description: participates in cell migration and adhesion, and in platelet aggregation  
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology  
 C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer  
 F;1-18/Domain: signal sequence #status predicted <SIG>  
 F;19-1170/Product: thrombospondin 1 #status predicted <MAT>  
 F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>  
 F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>  
 F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>  
 F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>  
 F;551-586/Domain: EGF homology <EGF1>  
 F;650-689/Domain: EGF homology <EGF2>  
 F;926-928/Region: cell attachment (R-G-D) motif  
 F;171-232/Disulfide bonds: #status predicted  
 F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F;270,274/Disulfide bonds: interchain #status predicted  
 F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted  
 F;1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 5.8%; Score 268.5; DB 1; Length 1170;  
 Best Local Similarity 32.9%; Pred. No. 4.1e-11;  
 Matches 51; Conservative 24; Mismatches 61; Indels 19; Gaps 4;

```

Qy      182 VVRQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGGWSTWTEWSVCSASC 233
      : :: | | : | | :: | : | : ||| | : || || : |
Db      399 IQQRGRSCDSLNNRCEGSSVQTRTCHIQECDKRFKQ-----DGGWSHWSPWSSCSVTC 451

Qy      234 GRGWQKRSRSC TNPAPLNGGAFCEGQNVQKTAC-ATLCPVDGSWSPWSKWSACGLDC--- 289
      | | | | | : | : | | | : : | | | : | | | | | : |
Db      452 GDGVITRIRLCNSPSPQMNGKPCEGEARETKACKKDACPINGGWGPWSPWDICSVTCGGG 511

Qy      290 THWRSRECS DPAPRNGGEECQGTDL DTRNCTSDLC 324
      ||| | : | | : | : | | : : | | |
Db      512 VQKRSRLCNNPTFQFGGKDCVGDVTENQICNKQDC 546
  
```

# RESULT 13

A39804

thrombospondin precursor - chicken

C;Species: Gallus gallus (chicken)

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004

C;Accession: A39804

R;Lawler, J.; Duquette, M.; Ferro, P.

J. Biol. Chem. 266, 8039-8043, 1991

A;Title: Cloning and sequencing of chicken thrombospondin.  
 A;Reference number: A39804; MUID:91217026; PMID:2022631  
 A;Accession: A39804  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-1178 <LAW>  
 A;Cross-references: UNIPROT:P35440; GB:M60853; NID:g212763; PIDN:AAA51437.1;  
 PID:g212764  
 C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat  
 homology; von Willebrand factor type C repeat homology  
 F;325-383/Domain: von Willebrand factor type C repeat homology <VWC>  
 F;386-437/Domain: thrombospondin type 1 repeat homology <THR1>  
 F;442-498/Domain: thrombospondin type 1 repeat homology <THR2>  
 F;499-555/Domain: thrombospondin type 1 repeat homology <THR3>  
 F;658-697/Domain: EGF homology <EGF>

Query Match 5.6%; Score 263; DB 1; Length 1178;  
 Best Local Similarity 36.2%; Pred. No. 1e-10;  
 Matches 58; Conservative 16; Mismatches 70; Indels 16; Gaps 5;

```

Qy      185 QARLADTANYTCVAKNIVARRRS-ASA AVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRS 243
      : | | | | : | | | | : : |||| | : || || : || | | |
Db      410 RGRSCDVTRSACTGPHIQTRMCSFKKCDHRIRQDGGWSHWSPWSSCSVTCGVGNITRIRL 469

Qy      244 CTNPAPLNGGAFCEGQNVQKTACATL-CPVDGSWSPWSKWSACGLDC---THWRSRECS 299
      | : | | || | | : | | | | | | | | | : | | | | :
Db      470 CNSPIQMGGKNCVGNGRETEKCEKAPCPVNGQWGPWPWSACTVTCGGGIRERSRLCNS 529

Qy      300 PAPRNGGEECQGTDLDT-----RNCTSDLCVHSASGP 331
      | | : || : | | | | | | | : | | | : : |
Db      530 PEPQYGGKPCVG---DTKQHDMCNKRDCPIDGCLSNPCFP 566
  
```

#### RESULT 14

S29126

properdin precursor [validated] - human

N;Alternate names: factor P

C;Species: Homo sapiens (man)

C;Date: 17-Nov-2000 #sequence\_revision 17-Nov-2000 #text\_change 09-Jul-2004

C;Accession: S29126; S16150; A05319; T45112; T45113

R;Nolan, K.F.; Kaluz, S.; Higgins, J.M.G.; Goundis, D.; Reid, K.B.M.

Biochem. J. 287, 291-297, 1992

A;Title: Characterization of the human properdin gene.

A;Reference number: S29126; MUID:93038568; PMID:1417780

A;Accession: S29126

A;Molecule type: DNA

A;Residues: 1-469 <NOL1>

A;Cross-references: UNIPROT:P27918; EMBL:X70872; NID:g35679; PIDN:CAA50220.1;  
 PID:g35680

R;Nolan, K.F.; Schwaeble, W.; Kaluz, S.; Dierich, M.P.; Reid, K.B.M.

Eur. J. Immunol. 21, 771-776, 1991

A;Title: Molecular cloning of the cDNA coding for properdin, a positive  
 regulator of the alternative pathway of human complement.

A;Reference number: S16150; MUID:91184288; PMID:2009915

A;Accession: S16150

A;Molecule type: mRNA

A;Residues: 1-456,'R',458-469 <NOL2>

A;Cross-references: EMBL:X57748  
 R;Reid, K.B.M.; Gagnon, J.  
 Mol. Immunol. 18, 949-959, 1981  
 A;Reference number: A05319; MUID:82195224; PMID:7341961  
 A;Accession: A05319  
 A;Molecule type: protein  
 A;Residues: 28-53,'Q',55-59,'G',61,'I',63;137-138,'P',140-141,'P',143-144,'X',146-148,'Y',150,'S',152,'Y',154-156,'XSXGXA';162-163,'E',165-172,'X',174-176,'X',178,'V',180;223-228,'X',230-232,'GX',235-238,'GH',241-245;248-251,'X',253-257,'P',259,'G',261,'XPP',265-266,'X',268-269;280-285,'X',287-290,'X',292,'H',294-300,'SXXX',305-307,'X',309-315,'K',317;333-341,343-357,'X',359-362,'EXE';393-404,'QK',407;421-427,'R',429-443,'TKV',447-448,'XX',451,'RX',454-455 <REI>  
 R;Westberg, J.; Nordin-Fredrikson, G.; Truedsson, L.; Sjöholm, A.G.; Uhlen, M. submitted to the EMBL Data Library, May 1997  
 A;Reference number: Z22914  
 A;Accession: T45112  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-54,'X',56-73,'X',75-99,'W',101-469 <WES1>  
 A;Cross-references: EMBL:AF005665; PIDN:AAB63280.1  
 A;Experimental source: genomic DNA from individual with properdin deficiency type II  
 A;Accession: T45113  
 A;Status: translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-60,'X',62-413,'D',415-452,'XX',455-469 <WE2>  
 A;Cross-references: EMBL:AF005666; PIDN:AAC51626.1  
 A;Experimental source: genomic DNA from individual with properdin deficiency type III  
 R;Hartmann, S.; Hofsteenge, J.  
 J. Biol. Chem. 275, 28569-28574, 2000  
 A;Title: Properdin, the positive regulator of complement, is highly C-mannosylated.  
 A;Reference number: A59360; MUID:20435812; PMID:10878002  
 A;Contents: annotation  
 A;Note: identification and location of C-mannosylation sites by mass-spectroscopy  
 C;Genetics:  
 A;Gene: GDB:PFC  
 A;Cross-references: GDB:120275; OMIM:312060  
 A;Map position: Xp11.3-Xp11.23  
 A;Introns: 26/1; 76/2; 135/1; 192/1; 256/1; 314/1; 378/1; 415/2  
 C;Complex: a mixture of homodimers, homotrimers and homotetramers  
 C;Function:  
 A;Description: protects C3 convertase (C3bBb) from rapid inactivation  
 A;Pathway: complement alternate pathway  
 C;Superfamily: human properdin precursor; thrombospondin type 1 repeat homology  
 C;Keywords: complement alternate pathway; glycoprotein; homodimer; homotetramer; homotrimer; plasma  
 F;1-27/Domain: signal sequence #status predicted <SIG>  
 F;28-469/Product: properdin #status experimental <MAT>  
 F;76-128/Domain: thrombospondin type 1 repeat homology <THR1>  
 F;135-191/Domain: thrombospondin type 1 repeat homology <THR2>  
 F;192-255/Domain: thrombospondin type 1 repeat homology <THR3>  
 F;256-313/Domain: thrombospondin type 1 repeat homology <THR4>  
 F;314-377/Domain: thrombospondin type 1 repeat homology <THR5>

F;378-440/Domain: thrombospondin type 1 repeat homology <THR6>  
F;83,86,139,142,145,196,199,260,263,321,324,382,385,388/Modified site: 2'-  
mannosyl-tryptophan (Trp) #status experimental  
F;428/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 5.2%; Score 243; DB 1; Length 469;  
Best Local Similarity 39.5%; Pred. No. 8.7e-10;  
Matches 45; Conservative 14; Mismatches 43; Indels 12; Gaps 4;

```
Qy      218 GGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACAT--LCPVDGS 275
          |||| | | || :| :| : | |:| :|| || | || : || | :|| |:
Db      137 GGWSGWPWEPCSVTCSKGRTRRRACNHPAPKCGG-HCPGQAQSEACDTQQVCPHGA 195

Qy      276 WSPWSKWSACGLDC-----THWRSRECSDPAP--RNGGEECQGTDLDTNCT 320
          |: | |: | | | ||:| | | : |: | | : | ||
Db      196 WATWGPWTPCSASCHGGPHEPKETRSRKCSAPEPSQKPPGKPCPLAYEQRRCT 249
```

RESULT 15

S05478

properdin - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 09-Jul-2004

C;Accession: S05478

R;Goundis, D.; Reid, K.B.M.

Nature 335, 82-85, 1988

A;Title: Properdin, the terminal complement components, thrombospondin and the circumsporozoite protein of malaria parasites contain similar sequence motifs.

A;Reference number: S05478; MUID:88318954; PMID:3045564

A;Accession: S05478

A;Molecule type: mRNA

A;Residues: 1-437 <GOU>

A;Cross-references: UNIPROT:P11680; EMBL:X12905; NID:g53786; PIDN:CAA31389.1;  
PID:g53787

C;Complex: a mixture of homodimers, homotrimers and homotetramers

C;Function:

A;Description: protects C3 convertase (C3bBb) from rapid inactivation

A;Pathway: complement alternate pathway

C;Superfamily: human properdin precursor; thrombospondin type 1 repeat homology

C;Keywords: complement alternate pathway; glycoprotein; homodimer; homotetramer;  
homotrimer; plasma

F;45-97/Domain: thrombospondin type 1 repeat homology <THR1>

F;104-160/Domain: thrombospondin type 1 repeat homology <THR2>

F;161-224/Domain: thrombospondin type 1 repeat homology <THR3>

F;225-282/Domain: thrombospondin type 1 repeat homology <THR4>

F;283-345/Domain: thrombospondin type 1 repeat homology <THR5>

F;346-408/Domain: thrombospondin type 1 repeat homology <THR6>

F;52,55,108,111,114,165,168,229,232,290,293,350,353,356/Modified site: 2'-  
mannosyl-tryptophan (Trp) #status predicted

F;366,396/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.9%; Score 229; DB 2; Length 437;  
Best Local Similarity 40.4%; Pred. No. 8.3e-09;  
Matches 46; Conservative 10; Mismatches 46; Indels 12; Gaps 4;

```
Qy      218 GGWSTWTEWSVCSASCGRGWQKRSRSTNPAPLNGGAFCEGQNVQKTACAT--LCPVDGS 275
          |||| | | || :| :| | | | |||| || | |: | || | || |:
```

Db 106 GGWSEWGPWGPCSVTCSKGTQIRQRVCDNPAPKCGG-HCPGEAQQSQACDTQKTCPTHGA 164  
Qy 276 WSPWSKWSACGLDC-----THWRSRECSDPAPRN--GGEECQGTDLDTNCT 320  
|: | || | ||| || ||| : |: | | : : |: |  
Db 165 WASWGPWSPRSGSCLGGAQEPKETRSRSCSAPAPSHQPPGKPCSGPAYEHKACS 218

Search completed: March 1, 2005, 09:07:18  
Job time : 36.0799 secs

OM protein - protein search, using sw model

Run on: March 1, 2005, 09:06:01 ; Search time 168.022 Seconds  
(without alignments)  
1704.439 Million cell updates/sec

Title: US-10-624-932-2\_COPY\_26\_898  
Perfect score: 4660  
Sequence: 1 QQSATVANPVPGANPDLLPH.....AVAGLGQPDAGLFTVSEAEC 873

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1385339 seqs, 328044528 residues

Total number of hits satisfying chosen parameters: 1385339

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep:\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4660	100.0	898	10	US-09-918-779-2	Sequence 2, Appli
2	4660	100.0	898	15	US-10-624-932-2	Sequence 2, Appli
3	4567.5	98.0	899	10	US-09-970-944-2	Sequence 2, Appli
4	4511	96.8	898	10	US-09-933-261-5	Sequence 5, Appli
5	4511	96.8	898	10	US-09-970-944-13	Sequence 13, Appl
6	4511	96.8	898	14	US-10-256-702-5	Sequence 5, Appli
7	4511	96.8	898	14	US-10-240-154-16	Sequence 16, Appl
8	4282	91.9	842	15	US-10-311-623-1	Sequence 1, Appli
9	2845	61.1	544	10	US-09-970-944-14	Sequence 14, Appl
10	2815.5	60.4	557	10	US-09-933-261-6	Sequence 6, Appli
11	2815.5	60.4	557	14	US-10-256-702-6	Sequence 6, Appli
12	2778.5	59.6	931	10	US-09-970-944-15	Sequence 15, Appl
13	2778.5	59.6	931	11	US-09-972-211-121	Sequence 121, App
14	2778.5	59.6	931	15	US-10-087-684-35	Sequence 35, Appl
15	2778.5	59.6	931	15	US-10-037-417-117	Sequence 117, App
16	2778.5	59.6	931	15	US-10-096-625-121	Sequence 121, App
17	2778.5	59.6	1010	15	US-10-218-779-35	Sequence 35, Appl
18	2754.5	59.1	931	10	US-09-970-944-16	Sequence 16, Appl
19	2754.5	59.1	931	11	US-09-972-211-125	Sequence 125, App
20	2754.5	59.1	931	15	US-10-096-625-125	Sequence 125, App
21	2751.5	59.0	931	15	US-10-037-417-120	Sequence 120, App
22	2747.5	59.0	931	10	US-09-970-944-17	Sequence 17, Appl
23	2747.5	59.0	931	11	US-09-972-211-122	Sequence 122, App
24	2747.5	59.0	931	15	US-10-087-684-36	Sequence 36, Appl
25	2747.5	59.0	931	15	US-10-218-779-36	Sequence 36, Appl
26	2747.5	59.0	931	15	US-10-037-417-118	Sequence 118, App
27	2747.5	59.0	931	15	US-10-037-417-119	Sequence 119, App
28	2747.5	59.0	931	15	US-10-096-625-122	Sequence 122, App
29	2560.5	54.9	943	10	US-09-933-261-7	Sequence 7, Appli
30	2560.5	54.9	943	14	US-10-256-702-7	Sequence 7, Appli
31	2560.5	54.9	945	11	US-09-972-211-124	Sequence 124, App
32	2560.5	54.9	945	15	US-10-087-684-34	Sequence 34, Appl
33	2560.5	54.9	945	15	US-10-218-779-34	Sequence 34, Appl
34	2560.5	54.9	945	15	US-10-037-417-121	Sequence 121, App
35	2560.5	54.9	945	15	US-10-096-625-124	Sequence 124, App
36	2557.5	54.9	945	11	US-09-972-211-123	Sequence 123, App
37	2557.5	54.9	945	15	US-10-087-684-33	Sequence 33, Appl
38	2557.5	54.9	945	15	US-10-218-779-33	Sequence 33, Appl
39	2557.5	54.9	945	15	US-10-096-625-123	Sequence 123, App
40	2546.5	54.6	933	15	US-10-087-684-2	Sequence 2, Appli
41	2546.5	54.6	933	15	US-10-087-684-4	Sequence 4, Appli
42	2546.5	54.6	933	15	US-10-218-779-2	Sequence 2, Appli
43	2546.5	54.6	933	15	US-10-218-779-4	Sequence 4, Appli
44	2546.5	54.6	945	15	US-10-037-417-38	Sequence 38, Appl
45	2541.5	54.5	945	14	US-10-028-072-146	Sequence 146, App

#### ALIGNMENTS

##### RESULT 1

US-09-918-779-2

; Sequence 2, Application US/09918779

; Publication No. US20030064369A1



```

; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook, John
; APPLICANT: Lepley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Smithson, Glennnda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/09/918,779
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/221,409
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,697
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/263,662
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: 60/281,645
; PRIOR FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-779-2

```

Query Match 100.0%; Score 4660; DB 10; Length 898;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	60
Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS	180
Db	146	RLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS	205
Qy	181	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP	300
Db	266	SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP	325
Qy	301	APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	360
Db	326	APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	385
Qy	361	KKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LSTTTTTYQGS LCPRQDGPS	420
Db	386	KKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LSTTTTTYQGS LCPRQDGPS	445
Qy	421	PKFQLTN GHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNMTYGT FNF	480
Db	446	PKFQLTN GHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNMTYGT FNF	505
Qy	481	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG	540
Db	506	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG	565
Qy	541	VLLTRPVILAMDHCGEPSPD SWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASAC Y	600
Db	566	VLLTRPVILAMDHCGEPSPD SWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASAC Y	625
Qy	601	VFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	660
Db	626	VFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	685
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSS LWKSKLLVSYQEIPFYHIWNGTQRYLH	720
Db	686	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSS LWKSKLLVSYQEIPFYHIWNGTQRYLH	745
Qy	721	CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	780
Db	746	CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	805
Qy	781	PSAFKIPFLIRQKI ISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA	840

Db 806 PSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA 865  
Qy 841 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 873  
|||||  
Db 866 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAEC 898

RESULT 2

US-10-624-932-2

; Sequence 2, Application US/10624932  
; Publication No. US20040096877A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier, Raymond  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Li, Li  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Grosse, William  
; APPLICANT: Alsobrook, John  
; APPLICANT: Lepley, Denise  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John  
; APPLICANT: Stone, David  
; APPLICANT: Smithson, Glennnda  
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-074 US  
; CURRENT APPLICATION NUMBER: US/10/624,932  
; CURRENT FILING DATE: 2003-07-21  
; PRIOR APPLICATION NUMBER: 09/918,779  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: 60/221,409  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: 60/222,840  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,752  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,762  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,770  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,769  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/225,146  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,392  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/225,470  
; PRIOR FILING DATE: 2000-08-15  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 898  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-624-932-2

Query Match 100.0%; Score 4660; DB 15; Length 898;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 873; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	60
Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSSGLPTMEVRINVSQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	180
Db	146	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDP SLDPNVYITREHS	205
Qy	181	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SRSCCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP	300
Db	266	SRSCCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP	325
Qy	301	APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLVLILVYCR	360
Db	326	APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLVLILVYCR	385
Qy	361	KKEGLDSDVADSSILTS GFQPVSIKPSKADNP HLLTIQPD LSTTTT TYQGS LCP RQD GPS	420
Db	386	KKEGLDSDVADSSILTS GFQPVSIKPSKADNP HLLTIQPD LSTTTT TYQGS LCP RQD GPS	445
Qy	421	PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FNF	480
Db	446	PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNM TYGT FNF	505
Qy	481	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGP PG	540
Db	506	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGP PG	565
Qy	541	VLLTRPVILAMDHCGEPSD SWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASAC Y	600
Db	566	VLLTRPVILAMDHCGEPSD SWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASAC Y	625
Qy	601	VFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	660
Db	626	VFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	685
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRYLH	720

Db 686 KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLH 745

Qy 721 CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG 780  
 |||

Db 746 CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG 805

Qy 781 PSAFKIPFLIRQKIISLDPCCRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEA 840  
 |||

Db 806 PSAFKIPFLIRQKIISLDPCCRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEA 865

Qy 841 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 873  
 |||

Db 866 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898

RESULT 3

US-09-970-944-2

; Sequence 2, Application US/09970944

; Publication No. US20030204052A1

; GENERAL INFORMATION:

; APPLICANT: Herrman, John L

; APPLICANT: Rastelli, Luca

; APPLICANT: Shimkets, Richard A

; TITLE OF INVENTION: No. US20030204052A1 Proteins and Nucleic Acids Encoding Same and

; TITLE OF INVENTION: Antibodies Directed Against these Proteins

; FILE REFERENCE: 21402-138

; CURRENT APPLICATION NUMBER: US/09/970,944

; CURRENT FILING DATE: 2002-05-02

; PRIOR APPLICATION NUMBER: 60/237,862

; PRIOR FILING DATE: 2000-10-04

; NUMBER OF SEQ ID NOS: 62

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 899

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-970-944-2

Query Match 98.0%; Score 4567.5; DB 10; Length 899;  
 Best Local Similarity 98.6%; Pred. No. 0;  
 Matches 863; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

Qy 1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60  
 |||

Db 26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy 61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 120  
 |||

Db 86 DHVIERSTDGSSGEPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 145

Qy 121 RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDPNVIYITREHS 180  
 |||

Db 146 RLRKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDPNVIYITREHS 205

Qy 181 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCASCGRWQKR 240  
 |||

Db	206	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS	299
		:	
Db	266	SRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSVDGSWSPWSKWSACGLDCTHWSRECS	325
Qy	300	PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLVLILVYC	359
Db	326	PAPRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLVLILVYC	385
Qy	360	RKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELLTIQPDLS-TTTTQYQGS	419
Db	386	RKKEGLSDSDVADSSILTSGFQPVSIKPSKADNPHELLTIQPDLS-TTTTQYQGS	444
Qy	420	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGT	479
Db	445	SPKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGT	504
Qy	480	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPP	539
Db	505	FLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPP	564
Qy	540	GVLLTRPVILAMDHCGEPSDWSLRLKKQSCGEQDVLHLGEEAPSHLYYCQLEASA	598
Db	565	GVLLTRPVILAMDHCGEPSDWSLRLKKQSCGEQDVLHLGEEAPSHLYYCQLEASA	624
Qy	599	CYVFTEQLGRFALVGEALSVAARKLKLFFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	658
Db	625	CYVFTEQLGRFALVGEALSVAARKLKLFFAPVACTSLEYNIRVYCLHDTHDALKEVVQ	684
Qy	659	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRY	718
Db	685	LEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRY	744
Qy	719	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	778
Db	745	LHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL	804
Qy	779	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLW	838
Db	805	VGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLW	864
Qy	839	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	873
Db	865	EARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	899

RESULT 4

US-09-933-261-5

; Sequence 5, Application US/09933261

; Publication No. US20030040046A1

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

```

; TITLE OF INVENTION: Netrin Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 BUSH STREET, SUITE 3200
; CITY: SAN FRANCISCO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/933,261
; FILING DATE: 20-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/808,982
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: OSMAN, RICHARD A
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030040046A1 Relevant
; TOPOLOGY: No. US20030040046A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-933-261-5

```

```

Query Match          96.8%; Score 4511; DB 10; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 838; Conservative 17; Mismatches 18; Indels 0; Gaps 0;

```

```

Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy      61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 120
        ||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      86 DHVIERSTDSSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 145

Qy     121 RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146 YLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS 205

Qy     181 LVVRQARLADTANYTCVAKNIVARRRSASAIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
        |||||||||||||||||||||||| ||||||||||||||||||||||||
Db     206 LVVRQARLADTANYTCVAKNIVARRRSTSAIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

```

Qy	241	SR SCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP	300
Db	266	SR SCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSSWSKWSACGLDCTHWSRECS DP	325
Qy	301	APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	360
		:       : :       : :                :	
Db	326	APRNGGEECRGADLDRNCTSDLC LHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCR	385
Qy	361	KKEGLSDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LSTTTTTYQGS LCPRQDGPS	420
Db	386	KKEGLSDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LSTTTTTYQGS LCSRQDGPS	445
Qy	421	PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAE EFVSRLSTQNYFRSLPRGTSNMTYGT FNF	480
		:           :	
Db	446	PKFQLSNGHLLSPLGSGRHTLHHSSPTSEAE DFVSRLSTQNYFRSLPRGTSNMAYGT FNF	505
Qy	481	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG	540
		:	
Db	506	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPV VSCGPPG	565
Qy	541	VLLTRPVILAMDHCGEPSDPSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACY	600
		:	
Db	566	VLLTRPVILAMDHCGEPSDPSWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGACY	625
Qy	601	VFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	660
		:	
Db	626	VFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	685
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSS LWSKLLVSYQEIPFYHIWNGTQRYLH	720
		:	
Db	686	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSS LWSKLLVSYQEIPFYHIWNGTQQYLH	745
Qy	721	CTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	780
		:       :     :	
Db	746	CTFTLERINASTSD LACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEGGVPALVG	805
Qy	781	PSAFKIPFLIRQKI ISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA	840
		:	
Db	806	PSAFKIPFLIRQKIIASLDPPCSR GADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA	865
Qy	841	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE C	873
Db	866	RHFPNGNLGQLAAAVAGLGQPDAGLFTVSEAE C	898

# RESULT 5

US-09-970-944-13

; Sequence 13, Application US/09970944

; Publication No. US20030204052A1

; GENERAL INFORMATION:

; APPLICANT: Herrman, John L

; APPLICANT: Rastelli, Luca

; APPLICANT: Shimkets, Richard A

; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and

; TITLE OF INVENTION: Antibodies Directed Against these Proteins



```
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-970-944-13
```

```
Query Match          96.8%; Score 4511; DB 10; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 838; Conservative 17; Mismatches 18; Indels 0; Gaps 0;
```

```
Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
        |||||||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy      61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEWVCQCAWSSSGTTKSQKAYIRIA 120
        |||||||
Db      86 DHVIERSTDSSSGLPTMEVRINVSRRQVEKVFGLLEEWVCQCAWSSSGTTKSQKAYIRIA 145

Qy     121 RLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS 180
        |||||||
Db     146 YLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS 205

Qy     181 LVVRQARLADTANYTCVAKNIVARRRSASAIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
        |||||||
Db     206 LVVRQARLADTANYTCVAKNIVARRRSTSAIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

Qy     241 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP 300
        |||||||
Db     266 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSSWSKWSACGLDCTHWSRECS DP 325

Qy     301 APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR 360
        |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db     326 APRNGGEECRGADLDRNCTSDLCVHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCR 385

Qy     361 KKEGLDSDVADSSILTSFGFQPVSIKPSKADNPHLLTIQPDLS TTTT TTYQGS LCP RQDGPS 420
        |||||||
Db     386 KKEGLDSDVADSSILTSFGFQPVSIKPSKADNPHLLTIQPDLS TTTT TTYQGS LCSRQDGPS 445

Qy     421 PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEVSRLSTQNYFRSLPRGTSNM TYGT FNF 480
        |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db     446 PKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEFEVSRLSTQNYFRSLPRGTSNMAYGT FNF 505

Qy     481 LGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG 540
        |||||||
Db     506 LGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPV VSCGPPG 565

Qy     541 VLLTRPVILAMDHCGEPSPD SWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACY 600
        |||||||
Db     566 VLLTRPVILAMDHCGEPSPD SWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEAGACY 625
```

Qy 601 VFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLH DTHDALKEVVQLE 660  
 |||:|||||  
 Db 626 VFTEQLGRFALVGEALSVAATKRLRLLLFAPVACTSLEYNIRVYCLH DTHDALKEVVQLE 685  
 Qy 661 KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKS KLLVSYQEIPFYHIWNGTQRYLH 720  
 |||:|||||  
 Db 686 KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKS KLLVSYQEIPFYHIWNGTQQYLH 745  
 Qy 721 CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG 780  
 |||:|||||:|||||:|||||:|||||  
 Db 746 CTFTLERINASTSDLACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEGGVPALVG 805  
 Qy 781 PSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA 840  
 |||:|||||:|||||:|||||:|||||  
 Db 806 PSAFKIPFLIRQKIIASLDPPCSR GADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA 865  
 Qy 841 RHF PNGNLSQLAAAVAGLGQPDAGLFTVSEAE C 873  
 |||  
 Db 866 RHF PNGNLGQLAAAVAGLGQPDAGLFTVSEAE C 898

RESULT 6

US-10-256-702-5

; Sequence 5, Application US/10256702

; Publication No. US20030059859A1

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/256,702

; FILING DATE: 27-Sep-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/933,261

; FILING DATE: 20-Aug-2001

; APPLICATION NUMBER: 08/808,982

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

```

; REFERENCE/DOCKET NUMBER: UC96-217
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 898 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030059859A1 Relevant
; TOPOLOGY: No. US20030059859A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-256-702-5

```

```

Query Match          96.8%; Score 4511; DB 14; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 838; Conservative 17; Mismatches 18; Indels 0; Gaps 0;

```

```

Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
      |||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy      61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEWVCQCVAWSSSGTTKSQKAYIRIA 120
      |||
Db      86 DHVIERSTDSSGLPTMEVRINVSRRQVEKVFGLLEEWVCQCVAWSSSGTTKSQKAYIRIA 145

Qy     121 RLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS 180
      |||
Db     146 YLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS 205

Qy     181 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
      |||
Db     206 LVVRQARLADTANYTCVAKNIVARRRSTSAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 265

Qy     241 SRSC TNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP 300
      |||
Db     266 SRSC TNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSSWSKWSACGLDCTHWSRECS DP 325

Qy     301 APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLLVILVYCR 360
      |||
Db     326 APRNGGEECRGADLDRNCTSDLCVHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCR 385

Qy     361 KKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPDLS TTTT TYQGS LCP RQDGPS 420
      |||
Db     386 KKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPDLS TTTT TYQGS LCSRQDGPS 445

Qy     421 PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNM TYGTFNF 480
      |||
Db     446 PKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMAYGTFNF 505

Qy     481 LGGR LMI PNTGISLLI PPDAI PRGKIYEIYLT LHKPEDVRLPLAGCQTLLS PIVSCGPPG 540
      |||
Db     506 LGGR LMI PNTGISLLI PPDAI PRGKIYEIYLT LHKPEDVRLPLAGCQTLLS PVVSCGPPG 565

Qy     541 VLLTRPVILAMDHCGEPS PDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACY 600
      |||
Db     566 VLLTRPVILAMDHCGEPS PDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEAGACY 625

```

Qy	601	VFTEQLGRFALVGEALSVA	AAKRLKLLLFAPVACTS	SLEYNIRVYCLHDTHDALKEVVQLE	660
Db	626	VFTEQLGRFALVGEALSVA	ATKRLRLLLFAPVACTS	SLEYNIRVYCLHDTHDALKEVVQLE	685
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRLSIH	VPSSLWKSLLVSYQEIPFYHIWNGTQRYLH	720	
Db	686	KQLGGQLIQEPRVLHFKDSYHNLRLSIH	VPSSLWKSLLVSYQEIPFYHIWNGTQQYLH	745	
Qy	721	CTFTLERVSPSTSDLACKLWVWQVEGDGQ	SFSINFNITKDTRFAELLALESEAGVPALVG	780	
Db	746	CTFTLERINASTSDLACKVWVWQVEGDGQ	SFNINFNITKDTRFAELLALESEGGVPALVG	805	
Qy	781	PSAFKIPFLIRQKIISSLDPPCRRGADWRT	LAQKLHLDHLSFFASKPSPTAMILNLWEA	840	
Db	806	PSAFKIPFLIRQKIIASLDPPCSR	GADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEA	865	
Qy	841	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSE	AEC	873	
Db	866	RHFPNGNLGQLAAAVAGLGQPDAGLFTVSE	AEC	898	

# RESULT 7

US-10-240-154-16

```

; Sequence 16, Application US/10240154
; Publication No. US20030175741A1
; GENERAL INFORMATION:
; APPLICANT: Cochran et al.
; TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES
; FILE REFERENCE: CKFW-P01-006
; CURRENT APPLICATION NUMBER: US/10/240,154
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: PCT/GB01/01486
; PRIOR FILING DATE: 2001-04-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 16
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Rattus sp.

```

US-10-240-154-16

```

Query Match          96.8%; Score 4511; DB 14; Length 898;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 838; Conservative 17; Mismatches 18; Indels 0; Gaps 0;

```

Qy	1	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	VLLVCKAVPATQIFFKCNGEWVRQV	60
Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKP	VLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVS	RQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDSSSGLPTMEVRINVS	RQQVEKVFGLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCR	PPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS	180
Db	146	YLRKNFEQEPLAKEVSLEQGIVLPCR	PPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS	205

Qy	181	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVRQARLADTANYTCVAKNIVARRRSTSAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SR SCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECS DP	300
Db	266	SR SCTNPAPLNGGAFCEGQNVQKTACATLCPVDGSWSWSKWSACGLDCTHWSRECS DP	325
Qy	301	APRNGGEECQGTDL DTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	360
		:       : :       : :           :	
Db	326	APRNGGEECRGADL DTRNCTSDLC LHTASCPEDVALYIGLVAVAVCLFLLLLALGLIYCR	385
Qy	361	KKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LSTTTTTTYQGS LCP RQDGPS	420
Db	386	KKEGLDSDVADSSILTS GFQPVSIKPSKADNPHLLTIQPD LSTTTTTTYQGS LCSRQDGPS	445
Qy	421	PKFQLTNGHLLSPLGG RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNM TYGTFNF	480
		:       :       :       :	
Db	446	PKFQLSNGHLLSPLGSG RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMAYGTFNF	505
Qy	481	LGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG	540
		:	
Db	506	LGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPV VSCGPPG	565
Qy	541	VLLTRPVILAMDHCGEPS PDSWSLRLKKQSCG SWEDVLHLGEEAPSHLYYCQLEASACY	600
		:	
Db	566	VLLTRPVILAMDHCGEPS PDSWSLRLKKQSCG SWEDVLHLGEESPSHLYYCQLEAGACY	625
Qy	601	VFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	660
		:	
Db	626	VFTEQLGRFALVGEALSVA ATKRLRLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	685
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRYLH	720
		:	
Db	686	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQQYLH	745
Qy	721	CTFTLERVSPSTSD LACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	780
		:       :       :	
Db	746	CTFTLERINASTSD LACKVWVWQVEGDGQSFNINFNITKDTRFAELLALESEGGVPALVG	805
Qy	781	PSAFKIPFLIRQKI ISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA	840
		:	
Db	806	PSAFKIPFLIRQKIIASLDPPCSR GADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEA	865
Qy	841	RHF PNGNLSQLAAAVAGLGQPDAGLFTVSEAEC	873
Db	866	RHF PNGNLGQLAAAVAGLGQPDAGLFTVSEAEC	898

# RESULT 8

US-10-311-623-1

; Sequence 1, Application US/10311623

; Publication No. US20040023244A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.

; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.

```
; APPLICANT: YUE, Henry; NGUYEN, Danniel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda
; APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.
; APPLICANT: DUGGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CD1
US-10-311-623-1
```

```
Query Match          91.9%; Score 4282; DB 15; Length 842;
Best Local Similarity 93.4%; Pred. No. 0;
Matches 815; Conservative 1; Mismatches 1; Indels 56; Gaps 1;
```

```
Qy      1 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 60
      |||
Db      26 QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV 85

Qy      61 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 120
      |||
Db      86 DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA 145

Qy     121 RLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS 180
      |||
Db     146 YLRKNFEQEPLAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHS 205

Qy     181 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR 240
      |||
Db     206 LVVRQARLADTANYTCVAKNIVARRRSASAAVIVY----- 240

Qy     241 SRSCTNPAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECS DP 300
      |||
Db     241 -----VDGWSWPWSKWSACGLDCTHWSRECS DP 269
```

Qy	301	APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	360
Db	270	APRNGGEECQGTDLDRNCTSDLCVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	329
Qy	361	KKEGLSDVADSSILTSQFQVSIKPSKADNPHLLTIQPDLSSTTTTYQGSICPRQDGPS	420
Db	330	KKEGLSDVADSSILTSQFQVSIKPSKADNPHLLTIQPDLSSTTTTYQGSICPRQDGPS	389
Qy	421	PKFQLTNHLLSPLGGGRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNF	480
Db	390	PKFQLTNHLLSPLGGGRHTLHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTNF	449
Qy	481	LGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPPG	540
Db	450	LGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLSPIVSCGPPG	509
Qy	541	VLLTRPVILAMDHCGEPSDWSLRLKKQSCGSEWEDVLHLGEEAPSHLYYCQLEASAC	600
Db	510	VLLTRPVILAMDHCGEPSDWSLRLKKQSCGSEWEDVLHLGEEAPSHLYYCQLEASAC	569
Qy	601	VFTEQLGRFALVGEALSVAARKLKLFFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	660
Db	570	VFTEQLGRFALVGEALSVAARKLKLFFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE	629
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRYLH	720
Db	630	KQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRYLH	689
Qy	721	CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	780
Db	690	CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG	749
Qy	781	PSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEA	840
Db	750	PSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEA	809
Qy	841	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	873
Db	810	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	842

RESULT 9

US-09-970-944-14

; Sequence 14, Application US/09970944

; Publication No. US20030204052A1

; GENERAL INFORMATION:

; APPLICANT: Herrman, John L

; APPLICANT: Rastelli, Luca

; APPLICANT: Shimkets, Richard A

; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same and

; TITLE OF INVENTION: Antibodies Directed Against these Proteins

; FILE REFERENCE: 21402-138

; CURRENT APPLICATION NUMBER: US/09/970,944

; CURRENT FILING DATE: 2002-05-02

; PRIOR APPLICATION NUMBER: 60/237,862

; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 544  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-970-944-14

Query Match 61.1%; Score 2845; DB 10; Length 544;  
Best Local Similarity 100.0%; Pred. No. 2.7e-227;  
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	333	DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSQGFQPVSIKPSKADNP	392
Db	4	DVALYVGLIAVAVCLVLLLLVLILVYCRKKEGLSDVADSSILTSQGFQPVSIKPSKADNP	63
Qy	393	HLLTIQPDLSSTTTTTYQGSCLPRQDGSPKQFQLTNGHLLSPLGGGRHTLHHSSPTSEAAE	452
Db	64	HLLTIQPDLSSTTTTTYQGSCLPRQDGSPKQFQLTNGHLLSPLGGGRHTLHHSSPTSEAAE	123
Qy	453	FVSRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT	512
Db	124	FVSRLSTQNYFRSLPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLT	183
Qy	513	LHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWLRLLKKQSCE	572
Db	184	LHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDSSWLRLLKKQSCE	243
Qy	573	GSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPV	632
Db	244	GSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAAPKRLKLLLFAPV	303
Qy	633	ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPS	692
Db	304	ACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPS	363
Qy	693	SLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFS	752
Db	364	SLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFS	423
Qy	753	INFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLA	812
Db	424	INFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLA	483
Qy	813	QKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAE	872
Db	484	QKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAGLGQPDAGLFTVSEAE	543
Qy	873	C 873	
Db	544	C 544	

RESULT 10  
US-09-933-261-6  
; Sequence 6, Application US/09933261





Qy	438	RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFTNFLGGRLMIPNTGISLLIP	497
Db	121	RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGTFTNFLGGRLMIPNTGISLLIP	180
Qy	498	PDAIPRGKIYEIYLTTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE	557
Db	181	PDAIPRGKIYEIYLTTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE	240
Qy	558	SPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	617
Db	241	SPDSWSLALKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALS	300
Qy	618	VAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFK	677
Db	301	VAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHLX	360
Qy	678	DSYHNLRLSIHDPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDIAC	737
Db	361	DSYHNLXLSXHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDIAC	420
Qy	738	KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS	797
Db	421	KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIIS	480
Qy	798	LDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAG	857
Db	481	LDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAVAG	540
Qy	858	LGQPDAGLFT-VSEAE	873
		:	
Db	541	TXPAGRWLLSQCEAE	557

RESULT 11

US-10-256-702-6

; Sequence 6, Application US/10256702

; Publication No. US20030059859A1

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30



Qy 738 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS 797  
 |||  
 Db 421 KLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISS 480  
 Qy 798 LDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG 857  
 |||  
 Db 481 LDPPCRRGADWRTLAQKLHLDShLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAG 540  
 Qy 858 LGQPDAGLFT-VSEAE 873  
 | : |||||  
 Db 541 TXPAGRWLLSQCSEAE 557

RESULT 12

US-09-970-944-15

; Sequence 15, Application US/09970944

; Publication No. US20030204052A1

; GENERAL INFORMATION:

; APPLICANT: Herrman, John L

; APPLICANT: Rastelli, Luca

; APPLICANT: Shimkets, Richard A

; TITLE OF INVENTION: No. US20030204052A1e1 Proteins and Nucleic Acids Encoding Same and

; TITLE OF INVENTION: Antibodies Directed Against these Proteins

; FILE REFERENCE: 21402-138

; CURRENT APPLICATION NUMBER: US/09/970,944

; CURRENT FILING DATE: 2002-05-02

; PRIOR APPLICATION NUMBER: 60/237,862

; PRIOR FILING DATE: 2000-10-04

; NUMBER OF SEQ ID NOS: 62

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 15

; LENGTH: 931

; TYPE: PRT

; ORGANISM: Caenorhabditis elegans

US-09-970-944-15

Query Match 59.6%; Score 2778.5; DB 10; Length 931;  
 Best Local Similarity 58.2%; Pred. No. 2e-221;  
 Matches 513; Conservative 150; Mismatches 195; Indels 23; Gaps 7;

Qy 11 PGANPDLLPHFLVEPEDVYIVKKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG 70  
 | | : |||||: ||| : ||||| | ||| |||||: ||| ||| | |||: |  
 Db 54 PSDPPEPLPHFLIEPEEAYIVKKNKPVNLYCKASPATQIYFKCNSEWVHQKDHVVDERVDE 113  
 Qy 71 SSSLPTMEVRINVSRRQVEKVFGLLEEWQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP 130  
 : ||| || | : |||||: ||| | : ||||| |||||: |||||: ||| ||| |||||  
 Db 114 TSGLIVREVSIEISRQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYLRKTFEQEP 173  
 Qy 131 LAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLAD 190  
 | ||||| : ||| ||||| |||||: |||: ||| | ||| : |||: |||||: |  
 Db 174 LGKEVSLEQEVLLQCRPEGIPVAEVEWLRNEDIIDPAEDRNFYITIDHNLIIKQARLSD 233  
 Qy 191 TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSTCTNPAPL 250  
 ||||| |||||: ||| : ||||| |||||: |||||: |||: |||: |||||  
 Db 234 TANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRTRTCTNPAPL 293



; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Grosse, William M  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Spytek, Kimberly Ann  
; APPLICANT: Li, Li  
; APPLICANT: Edinger, Shlomit  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John R  
; APPLICANT: Gunther, Erik  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Stone, David J  
; APPLICANT: Smithson, Glennnda  
; APPLICANT: Szekeres Jr, Edward S  
; TITLE OF INVENTION: No. US20040048245A1el Human Proteins, Polynucleotides  
Encoding Them And  
; TITLE OF INVENTION: Methods Of Using The Same  
; FILE REFERENCE: 21402-141  
; CURRENT APPLICATION NUMBER: US/09/972,211  
; CURRENT FILING DATE: 2001-10-05  
; PRIOR APPLICATION NUMBER: 60/238,325  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,323  
; PRIOR FILING DATE: 2000-10-05  
; PRIOR APPLICATION NUMBER: 60/238,400  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,397  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,401  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,379  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,402  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 30/238,384  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,373  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,372  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,383  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/238,382  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: 60/275,892  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/296,860  
; PRIOR FILING DATE: 2001-06-08  
; NUMBER OF SEQ ID NOS: 198  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 121  
; LENGTH: 931  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-972-211-121

Query Match 59.6%; Score 2778.5; DB 11; Length 931;  
 Best Local Similarity 58.2%; Pred. No. 2e-221;  
 Matches 513; Conservative 150; Mismatches 195; Indels 23; Gaps 7;

Qy	11	PGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG	70
		:      : :  :                  :          :	
Db	54	PSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQKDHVVDERUDE	113
Qy	71	SSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP	130
		:         :   : :    :   : : : : : : : :	
Db	114	TSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYLRKTFEQEP	173
Qy	131	LAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLAD	190
		:  :   : : : : : : : :         : : : : : :	
Db	174	LGKEVSLEQEVLLQCRPPGEGIPVAEVEWLRNEDIIDPAEDRNFYITIDHNLIIKQARLSD	233
Qy	191	TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPL	250
		: :  :  :   : : : : : : : : : : : : : : : :	
Db	234	TANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRTRTCTNPAPL	293
Qy	251	NGGAFCEGQNVQKTACATLCPVDGSWSWWSKWSACGLDCTHWSRECSDPAPRNGGEECQ	310
		: :             :        :        :   : : :	
Db	294	NGGAFCEGQSVQKIACTTLCVPDGRWTSWSKWSTCGTECTHWRRECTAPAPKNGGKDCD	353
Qy	311	GTDLDTNRCTSDLCVHSASGPEDVALYVGL-IAVAVCLVLLLVLILVYCRKKEGLDSDV	369
		:      : :  :   : :        : :  :    : :  :	
Db	354	GLVLQSKNCTDGLCMQAAPSDDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESDI	413
Qy	370	ADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSTTTTTYQGSCLPRQDGSPKFLTNHGH	429
		: :  :     :   : :  :  :     :	
Db	414	IDSSALNGGFQPVNIKAARQD---LLAVPPDLTSAAAMYRGPVYALHD-VSDKIPMTNSP	469
Qy	430	LLSPLGGGRHTLHHSS----PTSEAEFVSRLS---TQNYF-----RSLPRGT--S	471
		:    : :      :          :	
Db	470	ILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQSLLENEALNLKNQSLARQTDPS	529
Qy	472	NMTYGTGFNGLGRLMIPNTGISLLIPDAIPRGIYEIYLTLHKPEDVRLPLAGCQTLLS	531
		:        : : : :     : : : : : : : :    :    :	
Db	530	CTAFGTFNLSLGGHLIIPNSGVSLIPAGAIPOGRVYEMYVTVHRKENMRPPMEDSQTLLT	589
Qy	532	PIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCGWSWEDVLHLGEEAPSHLYY	591
		:            :    :   :   :     :      : :    :	
Db	590	PVVSCGPPGALLTRPVILTLLHHCADPSTEDWKIQLKNQAVQGQWEDVVVVEENFTTPCY	649
Qy	592	CQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHD	651
		:    :       :   : :  :         :  :  :   : :	
Db	650	IQLDAEACHILTENLSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQD	709
Qy	652	ALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWWSKLLVSYQEIPFYHI	711
		: : : : : : :             :	
Db	710	ALKEVLQLERQMGGQLLEEPKALHFKGSIHNLRLSIHDIAHSLWWSKLLAKYQEIPFYHI	769
Qy	712	WNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALES	771
		: :       : :  :  :          :    :  :   :  :	
Db	770	WSGSQRNLHCTFTLERLSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEPTGIDLPLDDP	829
Qy	772	EAGVPALVGPSAFKIPFLIRQKIISLDPPCRGRGADWRTLAQKLHLDShLSFFASKPSPT	831

```

      : : : ||||| || ||||: |||| | || ||| || ||:| | :|::||:| |||
Db      830 ASTITTVTGPSAFSIPPIRQKLCSSLDAPQTRGHDWRMLAHKLNLDRLNYFATKSSPT 889

Qy      832 AMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 872
      :||:||||:|:|:|||| ||| : :|: : : :| :
Db      890 GVILDLWEAQNFDPGNLSMLAAVLEEMGRHETVVSILAAEQ 930

```

RESULT 14

US-10-087-684-35

```

; Sequence 35, Application US/10087684
; Publication No. US20040029116A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: MacDougall, John R.
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; APPLICANT: Stone, David J.
; APPLICANT: Grosse, William M.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Cathereine E.
; APPLICANT: Casman, Stacie, J.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Li, Li
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Mishra, Vishnu
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Rastelli, Luca
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Guo, Xiaojia
; APPLICANT: Miller, Charles E.
; APPLICANT: Gangolli, Esha A.
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-214 CIP
; CURRENT APPLICATION NUMBER: US/10/087,684
; CURRENT FILING DATE: 2003-03-10
; PRIOR APPLICATION NUMBER: 60/253,834
; PRIOR FILING DATE: 2000-11-29
; PRIOR APPLICATION NUMBER: 60/250,926
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 60/264,180
; PRIOR FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/313,656
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/327,456
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 220
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 35
; LENGTH: 931

```



; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-087-684-35

Query Match 59.6%; Score 2778.5; DB 15; Length 931;  
Best Local Similarity 58.2%; Pred. No. 2e-221;  
Matches 513; Conservative 150; Mismatches 195; Indels 23; Gaps 7;

```
Qy      11 PGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG 70
      |  |: |||||:||||: ||||| | | | |||||:|||| | | | |||:  |
Db      54 PSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQKDHVVDERVDE 113

Qy      71 SSSLPTMEVRINVSRRQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP 130
      :||| | | | :|||||:|||| | :|||||:||||:||||:|||| | | | |||||
Db     114 TSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYLRKTFEQEP 173

Qy     131 LAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDPSPDPNVYITREHSLVVRQARLAD 190
      | ||||| :| | ||||| |||||:||||:| | | | | :|:|:|:|:|:|
Db     174 LGKEVSLEQEVLLQCRPPGIPVAEVEWLKNEDIIDPAEDRNFYITIDHNLIQKARLSD 233

Qy     191 TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPL 250
      |||||:|:| | :| |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     234 TANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRTCTNPAPL 293

Qy     251 NGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQ 310
      |||||:| | | ||||| | : |||| | | :||| | | | : |||:|:|:|
Db     294 NGGAFCEGQSVQKIACTTLCPWDGRWTSWSKWSTCGTECTHWRRRECTAPAPKNGGKDCD 353

Qy     311 GTDLDTNRCTSDLCVHSASGPEDVALYVGL-IAVAVCLVLLLVLILVYCRKKEGLDSDV 369
      | | :| | | | : | : |||||: | | | | : : | : | | : | | :
Db     354 GLVLQSKNCTDGLCMQAAPSDDDVALYGVIVIAVTVCLAITVVVALFVYRKNHRDFESDI 413

Qy     370 ADSSILTSGFQPVSIKPSKADNPHELLTIQPDLSSTTTTQYQSLCPRQDGSPKQLTNGH 429
      || | | |||||:| | : | | : |||: | : | : | | : | | : | |
Db     414 IDSSALNGGFQPVNIKAARQD---LLAVPPDLTSAAAMYRGPVYALHD-VSDKIPMTNSP 469

Qy     430 LLSPLGGGRHTLHHSS----PTSEAEFVSRLS---TQNYF-----RSLPRGT--S 471
      :| | | : :|| | | : | | | | | : | | : | | | | | | | | |
Db     470 ILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPQMTQSLLENEALNLKNQSLARQTDPS 529

Qy     472 NMTYGTFTNFLGGRLLMIPNTGISLLIPDPAIPRGIYEIYLTLHKPEDVRLPLAGCQTLLS 531
      :||| | | | :|||:|:| | | | |||:|:|:|:|:|:|:|:|:|:|:|:|
Db     530 CTAFGTFTNSLGGHLIIPNSGVSLIPAGAIPOGRVYEMYVTVHRKENMRPPMEDSQTLLT 589

Qy     532 PIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEEAPSHLYY 591
      |:| | | | | | | | | : | | : | | : | | | | : | | | : | | |
Db     590 PVVSCGPPGALLTRPVILT LHHCADPSTEDWKIQLKNQAVQGWEDVVVGEENFTTPCY 649

Qy     592 CQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHD 651
      ||:| ||:| | | | :|||:|:| | | | | | :| | :|:|:|:|:| | | |
Db     650 IQLDAEACHILTENLSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQD 709

Qy     652 ALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSLLVSYQEIPFYHI 711
      |||||:|:|:|:|:|:|:|:| | | | | | | | | | | | | | | | | |
Db     710 ALKEVLQLERQMGGQLLEPKALHFKGSIHNLRLSIHDIAHSLWKSLLAKYQEIPFYHI 769

Qy     712 WNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDQSFSINFNITKDTRFAELLALES 771
```



```

; PRIOR APPLICATION NUMBER: 60/303,231
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/305,060
; PRIOR FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: 60/318,405
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/318,700
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-037-417-117

```

```

Query Match          59.6%; Score 2778.5; DB 15; Length 931;
Best Local Similarity 58.2%; Pred. No. 2e-221;
Matches 513; Conservative 150; Mismatches 195; Indels 23; Gaps 7;

```

```

Qy      11 PGANPDLLPHFLVEPEDVYIVKNKPVLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG 70
      |  | : |||||:||||: ||||| | || |||||:|||| ||| | |||:: |
Db      54 PSDPPEPLPHFLIEPEEAYIVKNKPNVLYCKASPATQIYFKCNSEWVHQKDHVVDERVDE 113

Qy      71 SSSLPTMEVRINVSRRQVEKVFGLLEEYWCQCAWSSSGTTSQKAYIRIARLRKNFEQEP 130
      :||| || | :|||||:| || |:|||||:||||:||||:| || |||||
Db     114 TSGLIVREVSIEISRQQVEELFGPEDYWCQCAWSSAGTTKSRKAYVRIAYLRKTFEQEP 173

Qy     131 LAKEVSLEQGIVLPCRPPGEGIPPAEVEWLRNEDLVDPSPDPNVYITREHSLVVRQARLAD 190
      | ||||| :| ||||| |||||:||||:| | ||| :|:|:|:| |||:|
Db     174 LGKEVSLEQEVLLQCRPPGEGIPVAEVEWLRNEDIIDPAEDRNFYITIDHNLIIKQARLSD 233

Qy     191 TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPL 250
      |||||:|:| :| |||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db     234 TANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRTRCTNPAPL 293

Qy     251 NGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQ 310
      |||||:| || ||||| | : |||| || :||| |||: |||:|:|:|
Db     294 NGGAFCEGQSVQKIACCTLCPVDGRWTSWSKWSCTCGTECTHWRRECTAPAPKNGGKDCD 353

Qy     311 GTDLDRNCTSDLCVHSASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDV 369
      | | :| || ||: | : |||||: || || | : :| : | : ||:
Db     354 GLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVTVCLAITVVALFVYRKNHRDFESDI 413

Qy     370 ADSSILTSGFQPVSIKPSKADNPHELLTIQPDLTSTTTTYQGSCLPRQDGPSPKFQLTNGH 429
      ||| | |||||:| :| || : |||:| | : | | | :||
Db     414 IDSSALNGGFQPVNIKAARQD---LLAVPPDLTSAAAMYRGPVYALHD-VSDKIPMTNSP 469

Qy     430 LLSPLGGGRHTLHHSS----PTSEAEFVSRLS---TQNYF-----RSLPRGT--S 471
      :| || : :|| | : || | || : || : || | | |
Db     470 ILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLS PQMTQSLENEALNLKNQSLARQTDPS 529

Qy     472 NMTYGTFNFLGGRLMIPNTGISLLIPPDAIPRGIYIYLT LHKPEDVRLPLAGCQTLLS 531
      :||| || | :|||:|:| ||| |||:|:|:|:|:|:|:|:|:|:|:|
Db     530 CTAFGTFNSLGGHLIIPNSGVSLIPAGAI PQGRVYEMYVTVHRKENMRPPMEDSQTLLT 589

Qy     532 PIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWEDVLHLGEEAPSHLYY 591

```

Db	590	:                :     :     :   :       :       :       :     PVVSCGPPGALLTRPVILTLHHCADPSTEDWKIQLKNQAVQGQWEDVVVVGEENFTTPCY	649
Qy	592	CQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHD     :       :       :       :       :       :       :       :	651
Db	650	IQLDAEACHILTENLSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQD	709
Qy	652	ALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRLSIHDPSSLWWSKLLVSYQEIPFYHI         :       :       :       :       :       :       :       :	711
Db	710	ALKEVLQLERQMGGQLLEEPKALHFKGSIHNLRLSIHDIAHSLWWSKLLAKYQEIPFYHI	769
Qy	712	WNGTQRYLHCTFTLERVSPSTDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALES   :   :               :   :   :               :   :   :   :   :   :   :	771
Db	770	WSGSQRNHCTFTLERLSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEEP TGIDLPLLD	829
Qy	772	EAGVPALVGPSAFKIPFLIRQKIISLDP PCRRGADWRTLAQKLHLD SHLSFFASKPSPT : : :                   :                   :       :   :   :	831
Db	830	ASTITTVTGPSAFSIP LPIRQKLCSSLDAPQTRGHDWRMLAHKLNLD RYLN YFATKSSPT	889
Qy	832	AMILNLWEARHF PNGNLSQLAAAVAGLGQPDAGLFTVSEAE	872
Db	890	GVILDLWEAQNF PDGNLSMLAAVLEEMGRHETVVS LA AEGQ	930

Search completed: March 1, 2005, 09:51:29  
 Job time : 172.022 secs

OM protein - protein search, using sw model

Run on: March 1, 2005, 08:41:47 ; Search time 159.7 Seconds  
(without alignments)  
2799.282 Million cell updates/sec

Title: US-10-624-932-2\_COPY\_26\_898  
Perfect score: 4660  
Sequence: 1 QQSATVANPVPGANPDLLPH.....AVAGLGQPDAGLFTVSEAEC 873

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	% Query		Match	Length	DB	ID	Description
	Score	Match					
1	4558	97.8	898	1	UN5A_MOUSE	Q8k1s4	mus musculus
2	4511	96.8	898	1	UN5A_RAT	O08721	rattus norv
3	4279	91.8	842	1	UN5A_HUMAN	Q6zn44	homo sapien
4	2778.5	59.6	931	1	UN5C_MOUSE	O08747	mus musculus
5	2762.5	59.3	931	1	UN5C_RAT	Q761x5	rattus norv
6	2750.5	59.0	931	1	UN5C_CHICK	Q7t2z5	gallus gall
7	2747.5	59.0	931	1	UN5C_HUMAN	O95185	homo sapien
8	2639	56.6	943	1	UN5B_XENLA	Q8jgt4	xenopus lae
9	2563.5	55.0	945	1	UN5B_MOUSE	Q8k1s3	mus musculus
10	2560.5	54.9	945	1	UN5B_RAT	O08722	rattus norv
11	2541.5	54.5	945	1	UN5B_HUMAN	Q8izj1	homo sapien
12	2185.5	46.9	953	1	UN5D_HUMAN	Q6uxz4	homo sapien
13	2183	46.8	956	1	UN5D_MOUSE	Q8k1s2	mus musculus
14	1050.5	22.5	876	2	Q7PW78	Q7pw78	anopheles g
15	992	21.3	1072	1	UNC5_DROME	Q95tu8	drosophila

16	977	21.0	919	1	UNC5_CAEEL	Q26261	caenorhabdi
17	891.5	19.1	759	2	Q7PW77	Q7pw77	anopheles g
18	692	14.8	199	1	UNC5_PETMA	Q9pvd5	petromyzon
19	377.5	8.1	2673	2	Q96SC3	Q96sc3	homo sapien
20	377.5	8.1	5636	2	Q96RW7	Q96rw7	homo sapien
21	370.5	8.0	1244	2	Q69YJ3	Q69yj3	homo sapien
22	333	7.1	1388	2	Q7QKD0	Q7qkd0	anopheles g
23	318	6.8	325	2	Q8I1K1	Q8ilk1	drosophila
24	300	6.4	518	2	Q8IV45	Q8iv45	homo sapien
25	298.5	6.4	1584	1	BAI1_HUMAN	O14514	homo sapien
26	296.5	6.4	1172	1	TSP2_HUMAN	P35442	homo sapien
27	293	6.3	1074	1	SM5A_HUMAN	Q13591	homo sapien
28	293	6.3	1172	1	TSP2_MOUSE	Q03350	mus musculu
29	293	6.3	1172	2	Q7TMT3	Q7tmt3	mus musculu
30	293	6.3	1172	2	Q8CG21	Q8cg21	mus musculu
31	292	6.3	1582	2	Q8CGM0	Q8cgm0	mus musculu
32	291.5	6.3	1170	1	TSP2_BOVIN	Q95116	bos taurus
33	291	6.2	1077	1	SM5A_MOUSE	Q62217	mus musculu
34	288.5	6.2	1088	2	Q6PCK8	Q6pck8	xenopus lae
35	286	6.1	1081	2	Q9U631	Q9u631	drosophila
36	285	6.1	1091	2	Q7YU67	Q7yu67	drosophila
37	285	6.1	1093	2	Q9VTT0	Q9vtt0	drosophila
38	284	6.1	518	2	Q6R653	Q6r653	mus musculu
39	280	6.0	1173	1	TSP1_XENLA	P35448	xenopus lae
40	276	5.9	1093	1	SM5B_HUMAN	Q9p283	homo sapien
41	276	5.9	1151	2	Q6DD89	Q6dd89	homo sapien
42	276	5.9	1461	2	Q8MYA8	Q8mya8	caenorhabdi
43	275.5	5.9	632	2	Q6ZPQ8	Q6zpq8	mus musculu
44	275.5	5.9	1093	1	SM5B_MOUSE	Q60519	mus musculu
45	275.5	5.9	1122	2	Q7TT33	Q7tt33	mus musculu

#### ALIGNMENTS

##### RESULT 1

##### UN5A\_MOUSE

ID UN5A\_MOUSE STANDARD; PRT; 898 AA.  
AC Q8K1S4; Q6PEF7; Q80T71;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Netrin receptor UNC5A precursor (Unc-5 homolog A) (Unc-5 homolog 1).  
GN Name=Unc5a; Synonyms=Kiaa1976, Unc5h1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.  
RX MEDLINE=22239710; PubMed=12351186; DOI=10.1016/S0925-4773(02)00248-4;  
RA Engelkamp D.;  
RT "Cloning of three mouse unc-5 genes and their expression patterns at  
RT mid-gestation.";  
RL Mech. Dev. 118:191-197(2002).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 3).

RC TISSUE=Brain;  
 RX MEDLINE=22579291; PubMed=12693553;  
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,  
 RA Nakajima D., Nagase T., Ohara O., Koga H.;  
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:  
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous  
 RT cDNAs identified by screening of terminal sequences of cDNA clones  
 RT randomly sampled from size-fractionated libraries.";  
 RL DNA Res. 10:35-48(2003).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates  
 CC axon repulsion of neuronal growth cones in the developing nervous  
 CC system upon ligand binding. Axon repulsion in growth cones may be  
 CC caused by its association with DCC that may trigger signaling for  
 CC repulsion. It also acts as a dependence receptor required for  
 CC apoptosis induction when not associated with netrin ligand (By  
 CC similarity).  
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC. Interacts  
 CC with MAGED1. Interacts with PRKCABP, possibly mediating some  
 CC interaction with PKC (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. The interaction  
 CC with PRKCABP regulates its surface expression and leads to its  
 CC removal from surface of neurons and growth cones (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1;  
 CC IsoId=Q8K1S4-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q8K1S4-2; Sequence=VSP\_011697;  
 CC Note=No experimental confirmation available;  
 CC Name=3;  
 CC IsoId=Q8K1S4-3; Sequence=VSP\_011696;  
 CC Note=No experimental confirmation available;  
 CC -!- TISSUE SPECIFICITY: Restricted to central nervous system.

CC -!- DOMAIN: The ZU5 domain mediates the interaction with MAGED1, which  
 CC participates in the induction of apoptosis (By similarity).  
 CC -!- PTM: Phosphorylated by PKC in vitro. Phosphorylated on cytoplasmic  
 CC tyrosine residues (By similarity).  
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The  
 CC cleavage does not take place when the receptor is associated with  
 CC netrin ligand. Its cleavage by caspases is required to induce  
 CC apoptosis (By similarity).  
 CC -!- SIMILARITY: Belongs to the UNC-5 family.  
 CC -!- SIMILARITY: Contains 1 death domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC -!- SIMILARITY: Contains 1 ZU5 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ487852; CAD32250.1; -.  
 DR EMBL; AK122575; BAC65857.1; ALT\_INIT.  
 DR EMBL; BC058084; AAH58084.1; -.  
 DR HSSP; P07996; 1LSL.  
 DR MGD; MGI:894682; Unc5a.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR011029; DEATH\_like.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.  
 DR InterPro; IPR000906; ZU5.  
 DR Pfam; PF00531; Death; 1.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00090; TSP\_1; 2.  
 DR Pfam; PF00791; ZU5; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00409; IG; 1.  
 DR SMART; SM00209; TSP1; 2.  
 DR SMART; SM00218; ZU5; 1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; FALSE\_NEG.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS50092; TSP1; 2.  
 KW Alternative splicing; Apoptosis; Developmental protein;  
 KW Immunoglobulin domain; Phosphorylation; Receptor; Repeat; Signal;  
 KW Transmembrane.  
 FT SIGNAL 1 25 Potential.  
 FT CHAIN 26 898 Netrin receptor UNC5A.  
 FT DOMAIN 26 361 Extracellular (Potential).  
 FT TRANSMEM 362 382 Potential.  
 FT DOMAIN 383 898 Cytoplasmic (Potential).  
 FT DOMAIN 44 141 Ig-like.  
 FT DOMAIN 155 234 Ig-like C2-type.



FT	DOMAIN	242	296	TSP type-1 1.
FT	DOMAIN	298	350	TSP type-1 2.
FT	DOMAIN	495	598	ZU5.
FT	DOMAIN	817	897	Death.
FT	SITE	396	397	Cleavage (by caspase-3) (By similarity).
FT	SITE	661	679	Interaction with DCC (By similarity).
FT	DISULFID	65	124	By similarity.
FT	DISULFID	170	221	By similarity.
FT	CARBOHYD	107	107	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	218	218	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	343	343	N-linked (GlcNAc. . .) (Potential).
FT	VARSPLIC	1	790	Missing (in isoform 3).
FT				/FTId=VSP_011696.
FT	VARSPLIC	241	296	Missing (in isoform 2).
FT				/FTId=VSP_011697.
FT	CONFLICT	217	217	A -> P (in Ref. 3).
SQ	SEQUENCE	898 AA;	98856 MW;	59F04BA2E196C1DB CRC64;

Query Match 97.8%; Score 4558; DB 1; Length 898;

Best Local Similarity 96.7%; Pred. No. 0;

Matches 844; Conservative 19; Mismatches 10; Indels 0; Gaps 0;

Qy	1	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	60
Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS	180
Db	146	YLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS	205
Qy	181	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	265
Qy	241	SRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECSDF	300
Db	266	SRSCNPNAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECSDF	325
Qy	301	APRNGGEECQGTDLDTNRCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	360
Db	326	APRNGGEECRGADLDTRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	385
Qy	361	KKEGLDSDVADSSILTSFGFQPVSIKPSKADNPHLLTIQPDLSSTTTTYQGSCLPRQDGPS	420
Db	386	KKEGLDSDVADSSILTSFGFQPVSIKPSKADNPHLLTIQPDLSSTTTTYQGSCLPRQDGPS	445
Qy	421	PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFEVSRLSTQNYFRSLPRGTSNMTYGTNF	480
Db	446	PKFQLSNGHLLSPLGSGRHTLHHSSPTSEAEFEVSRLSTQNYFRSLPRGTSNMTYGTNF	505
Qy	481	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPG	540
Db	506	LGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPG	565

Qy 541 VLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACY 600  
 |||||:|||||  
 Db 566 VLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEESPSHLYYCQLEAGACY 625  
 Qy 601 VFTEQLGRFALVGEALSVAATAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE 660  
 |||||:|||||  
 Db 626 VFTEQLGRFALVGEALSVAATAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLE 685  
 Qy 661 KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLH 720  
 |||||:|||||  
 Db 686 KQLGGQLIQEPRVLHFKDSYHNLRLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLH 745  
 Qy 721 CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG 780  
 |||||:|||||:|||||:|||||:|||||  
 Db 746 CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVG 805  
 Qy 781 PSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKLHLDLHLSFFASKPSPTAMILNLWEA 840  
 |||||:|||||  
 Db 806 PSAFKIPFLIRQKIISLDPPCRRGADWRTLAQKLHLDLHLSFFASKPSPTAMILNLWEA 865  
 Qy 841 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 873  
 |||||  
 Db 866 RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 898

# RESULT 2

## UN5A\_RAT

ID UN5A\_RAT STANDARD; PRT; 898 AA.  
 AC 008721;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Netrin receptor UNC5A precursor (Unc-5 homolog A) (Unc-5 homolog 1).  
 GN Name=Unc5a; Synonyms=Unc5hl;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE  
 RP SPECIFICITY.  
 RC TISSUE=Ventral spinal cord;  
 RX MEDLINE=97271897; PubMed=9126742;  
 RA Leonardo E.D., Hinck L., Masu M., Keino-Masu K., Ackerman S.L.,  
 RA Tessier-Lavigne M.;  
 RT "Vertebrate homologues of C. elegans UNC-5 are candidate netrin  
 RT receptors.";  
 RL Nature 386:833-838(1997).  
 RN [2]  
 RP FUNCTION, AND INTERACTION WITH DCC.  
 RX PubMed=10399920;  
 RA Hong K., Hinck L., Nishiyama M., Poo M.-M., Tessier-Lavigne M.,  
 RA Stein E.;  
 RT "A ligand-gated association between cytoplasmic domains of UNC5 and  
 RT DCC family receptors converts netrin-induced growth cone attraction to  
 RT repulsion.";

RL Cell 97:927-941(1999).  
 RN [3]  
 RP TISSUE SPECIFICITY.  
 RX PubMed=11472849;  
 RA Barrett C., Guthrie S.;  
 RT "Expression patterns of the netrin receptor UNC5H1 among developing  
 RT motor neurons in the embryonic rat hindbrain.";  
 RL Mech. Dev. 106:163-166(2001).  
 RN [4]  
 RP FUNCTION.  
 RX PubMed=11387206; DOI=10.1093/emboj/20.11.2715;  
 RA Llambi F., Causeret F., Bloch-Gallego E., Mehlen P.;  
 RT "Netrin-1 acts as a survival factor via its receptors UNC5H and DCC.";  
 RL EMBO J. 20:2715-2722(2001).  
 RN [5]  
 RP FUNCTION, SUBCELLULAR LOCATION, AND INTERACTION WITH MAGED1.  
 RX PubMed=12598531; DOI=10.1074/jbc.M300415200;  
 RA Williams M.E., Strickland P., Watanabe K., Hinck L.;  
 RT "UNC5H1 induces apoptosis via its juxtamembrane region through an  
 RT interaction with NRAGE.";  
 RL J. Biol. Chem. 278:17483-17490(2003).  
 RN [6]  
 RP INTERACTION WITH PRKCABP, PHOSPHORYLATION, AND MUTAGENESIS OF  
 RP 896-ALA--CYS-898.  
 RX PubMed=14672991; DOI=23/36/11279;  
 RA Williams M.E., Wu S.C.-Y., McKenna W.L., Hinck L.;  
 RT "Surface expression of the netrin receptor UNC5H1 is regulated through  
 RT a protein kinase C-interacting protein/protein kinase-dependent  
 RT mechanism.";  
 RL J. Neurosci. 23:11279-11288(2003).  
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates  
 CC axon repulsion of neuronal growth cones in the developing nervous  
 CC system upon ligand binding. Axon repulsion in growth cones may be  
 CC caused by its association with DCC that may trigger signaling for  
 CC repulsion. It also acts as a dependence receptor required for  
 CC apoptosis induction when not associated with netrin ligand.  
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC. Interacts  
 CC with MAGED1. Interacts with PRKCABP, possibly mediating some  
 CC interaction with PKC.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. The interaction  
 CC with PRKCABP regulates its surface expression and leads to its  
 CC removal from surface of neurons and growth cones.  
 CC -!- TISSUE SPECIFICITY: Mainly expressed in regions of differentiating  
 CC neurons. Expressed at early stages of neural tube development in  
 CC the ventral spinal cord. In developing hindbrain, it colocalizes  
 CC with a number of cranial motor neuron subpopulations from  
 CC embryonic E11 to E14, while DCC is expressed by motor neurons at  
 CC E12. Also expressed in non-neural structures, such as the basal  
 CC plane of the hindbrain and midbrain, in the developing  
 CC hypothalamus, thalamus and in the pallidum.  
 CC -!- DOMAIN: The ZU5 domain mediates the interaction with MAGED1, which  
 CC participates in the induction of apoptosis.  
 CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By  
 CC similarity). Phosphorylated by PKC in vitro.  
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The  
 CC cleavage does not take place when the receptor is associated with  
 CC netrin ligand. Its cleavage by caspases is required to induce

```

CC      apoptosis.
CC      -!- SIMILARITY: Belongs to the UNC-5 family.
CC      -!- SIMILARITY: Contains 1 death domain.
CC      -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC      -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC      -!- SIMILARITY: Contains 2 TSP type-1 domains.
CC      -!- SIMILARITY: Contains 1 ZU5 domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U87305; AAB57678.1; -.
DR      HSSP; P07996; 1LSL.
DR      RGD; 621755; Unc5hl.
DR      InterPro; IPR000488; Death.
DR      InterPro; IPR011029; DEATH_like.
DR      InterPro; IPR003599; Ig.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      InterPro; IPR000906; ZU5.
DR      Pfam; PF00531; Death; 1.
DR      Pfam; PF00047; ig; 1.
DR      Pfam; PF00090; TSP_1; 2.
DR      Pfam; PF00791; ZU5; 1.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00005; DEATH; 1.
DR      SMART; SM00409; IG; 1.
DR      SMART; SM00209; TSP1; 2.
DR      SMART; SM00218; ZU5; 1.
DR      PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
DR      PROSITE; PS50835; IG_LIKE; 1.
DR      PROSITE; PS50092; TSP1; 2.
KW      Apoptosis; Developmental protein; Immunoglobulin domain;
KW      Phosphorylation; Receptor; Repeat; Signal; Transmembrane.
FT      SIGNAL          1      25      Potential.
FT      CHAIN           26     898      Netrin receptor UNC5A.
FT      DOMAIN          26     361      Extracellular (Potential).
FT      TRANSMEM        362     382      Potential.
FT      DOMAIN          383     898      Cytoplasmic (Potential).
FT      DOMAIN          44     141      Ig-like.
FT      DOMAIN          155     238      Ig-like C2-type.
FT      DOMAIN          242     296      TSP type-1 1.
FT      DOMAIN          298     350      TSP type-1 2.
FT      DOMAIN          495     598      ZU5.
FT      DOMAIN          817     897      Death.
FT      SITE            396     397      Cleavage (by caspase-3) (By similarity).
FT      SITE            661     679      Interaction with DCC (By similarity).
FT      DISULFID         65     124      By similarity.
FT      DISULFID        170     221      By similarity.
FT      CARBOHYD         107     107      N-linked (GlcNAc . . .) (Potential).
FT      CARBOHYD        218     218      N-linked (GlcNAc . . .) (Potential).

```





RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45(2004).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 302-842 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE OF 624-728 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21842142; PubMed=11853319;  
 RA Nagase T., Kikuno R., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XXII.  
 RT The complete sequences of 50 new cDNA clones which code for large  
 RT proteins.";  
 RL DNA Res. 8:319-327(2001).  
 RN [4]  
 RP INDUCTION.  
 RX PubMed=12598906; DOI=10.1038/ncb943;  
 RA Tanikawa C., Matsuda K., Fukuda S., Nakamura Y., Arakawa H.;  
 RT "p53RDL1 regulates of p53-dependent apoptosis.";  
 RL Nat. Cell Biol. 5:216-223(2003).  
 RN [5]  
 RP DOWN-REGULATION IN CANCER.  
 RX PubMed=12655055; DOI=10.1073/pnas.0738063100;  
 RA Thiebault K., Mazelin L., Pays L., Llambi F., Joly M.-O.,  
 RA Scoazec J.-Y., Saurin J.-C., Romeo G., Mehlen P.;  
 RT "The netrin-1 receptors UNC5H are putative tumor suppressors  
 RT controlling cell death commitment.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:4173-4178(2003).  
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates  
 CC axon repulsion of neuronal growth cones in the developing nervous  
 CC system upon ligand binding. Axon repulsion in growth cones may be  
 CC caused by its association with DCC that may trigger signaling for  
 CC repulsion. It also acts as a dependence receptor required for  
 CC apoptosis induction when not associated with netrin ligand.

CC    -!- SUBUNIT: Interacts with the cytoplasmic part of DCC. Interacts  
 CC       with MAGED1. Interacts with PRKCABP, possibly mediating some  
 CC       interaction with PKC (By similarity).  
 CC    -!- SUBCELLULAR LOCATION: Type I membrane protein. The interaction  
 CC       with PRKCABP regulates its surface expression and leads to its  
 CC       removal from surface of neurons and growth cones (By similarity).  
 CC    -!- ALTERNATIVE PRODUCTS:  
 CC       Event=Alternative splicing; Named isoforms=3;  
 CC       Name=1;  
 CC            IsoId=Q6ZN44-1; Sequence=Displayed;  
 CC            Note=No experimental confirmation available;  
 CC       Name=2;  
 CC            IsoId=Q6ZN44-2; Sequence=VSP\_011694, VSP\_011695;  
 CC            Note=No experimental confirmation available;  
 CC       Name=3;  
 CC            IsoId=Q6ZN44-3; Sequence=VSP\_011693;  
 CC            Note=No experimental confirmation available;  
 CC    -!- INDUCTION: By p53/TP53.  
 CC    -!- DOMAIN: The ZU5 domain mediates the interaction with MAGED1, which  
 CC       participates in the induction of apoptosis (By similarity).  
 CC    -!- PTM: Phosphorylated on cytoplasmic tyrosine residues.  
 CC       Phosphorylated by PKC in vitro (By similarity).  
 CC    -!- PTM: Proteolytically cleaved by caspases during apoptosis. The  
 CC       cleavage does not take place when the receptor is associated with  
 CC       netrin ligand. Its cleavage by caspases is required to induce  
 CC       apoptosis (By similarity).  
 CC    -!- MISCELLANEOUS: Down-regulated in multiple cancers including  
 CC       colorectal, breast, ovary, uterus, stomach, lung, or kidney  
 CC       cancers.  
 CC    -!- SIMILARITY: Belongs to the UNC-5 family.  
 CC    -!- SIMILARITY: Contains 1 death domain.  
 CC    -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC    -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC    -!- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC    -!- SIMILARITY: Contains 1 ZU5 domain.  
 CC    -!- CAUTION: Ref.3 sequence differs from that shown due to the  
 CC       presence of introns.

---

CC    This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC    the European Bioinformatics Institute. There are no restrictions on its  
 CC    use by non-profit institutions as long as its content is in no way  
 CC    modified and this statement is not removed. Usage by and for commercial  
 CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

---

DR    EMBL; AK131380; BAD18531.1; -.  
 DR    EMBL; BC009333; AAH09333.2; -.  
 DR    EMBL; BC033727; -; NOT\_ANNOTATED\_CDS.  
 DR    EMBL; AB075856; BAB85562.1; ALT\_SEQ.  
 DR    Genew; HGNC:12567; UNC5A.  
 DR    MIM; 607869; -.  
 DR    InterPro; IPR003599; Ig.  
 DR    InterPro; IPR007110; Ig-like.  
 DR    InterPro; IPR000884; TSP1.  
 DR    Pfam; PF00047; ig; 1.  
 DR    Pfam; PF00090; TSP\_1; 1.



DR SMART; SM00409; IG; 1.  
DR SMART; SM00209; TSP1; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; FALSE\_NEG.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
DR PROSITE; PS50092; TSP1; 1.  
KW Alternative splicing; Apoptosis; Developmental protein;  
KW Immunoglobulin domain; Phosphorylation; Receptor; Signal;  
KW Transmembrane.

FT	SIGNAL	1	25	Potential.
FT	CHAIN	26	842	Netrin receptor UNC5A.
FT	DOMAIN	1	306	Extracellular (Potential).
FT	TRANSMEM	307	327	Potential.
FT	DOMAIN	328	842	Cytoplasmic (Potential).
FT	DOMAIN	44	141	Ig-like.
FT	DOMAIN	155	234	Ig-like C2-type.
FT	DOMAIN	242	294	TSP type-1.
FT	DOMAIN	439	542	ZU5.
FT	DOMAIN	761	841	Death.
FT	SITE	340	341	Cleavage (by caspase-3) (By similarity).
FT	SITE	605	623	Interaction with DCC (By similarity).
FT	DISULFID	65	124	By similarity.
FT	DISULFID	170	221	By similarity.
FT	CARBOHYD	107	107	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	218	218	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	287	287	N-linked (GlcNAc. . .) (Potential).
FT	VARSP LIC	1	97	MAVRPGLWPALLGIVLAAWLRGSGAQQSATVANPVPGANPD
FT				LLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWV
FT				RQVDHVIERSTDGSN -> MAGTSERSLISSISQPKAIECF
FT				EVKKKAFLTHGRYHGSGATPPKTKDPKPETFCGQT (in
FT				isoform 3).
FT				/FTId=VSP_011693.
FT	VARSP LIC	296	301	TASGPE -> SESSLP (in isoform 2).
FT				/FTId=VSP_011694.
FT	VARSP LIC	302	842	Missing (in isoform 2).
FT				/FTId=VSP_011695.
SQ	SEQUENCE	842 AA;	92958 MW;	3DFADCF973131849 CRC64;

Query Match 91.8%; Score 4279; DB 1; Length 842;  
Best Local Similarity 93.2%; Pred. No. 1.4e-313;  
Matches 814; Conservative 2; Mismatches 1; Indels 56; Gaps 1;

Qy	1	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	60
Db	26	QQSATVANPVPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQV	85
Qy	61	DHVIERSTDGSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	120
Db	86	DHVIERSTDGSNGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIA	145
Qy	121	RLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS	180
Db	146	YLRKNFEQEPLAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHS	205
Qy	181	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKR	240
Db	206	LVVRQARLADTANYTCVAKNIVARRRSASAAVIVY-----	240

Qy	241	SRSCTNPAPLNGGAFCEGQNVQKTACATLCVPDGSWSPWSKWSACGLDCTHWSRECS	300
Db	241	-----VDGSWSEWSPWSKWSACGLDCTHWSRECS	269
Qy	301	APRNGGEECQGTDLDRNCTSDLCVHSASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	360
		:	
Db	270	APRNGGEECQGTDLDRNCTSDLCVHTASGPEDVALYVGLIAVAVCLVLLLLVLILVYCR	329
Qy	361	KKEGLDSDVADSSILTSFGQPVSIKPSKADNPHELLTIQPDLSSTTTTYQGS	420
Db	330	KKEGLDSDVADSSILTSFGQPVSIKPSKADNPHELLTIQPDLSSTTTTYQGS	389
Qy	421	PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGT	480
Db	390	PKFQLTNGHLLSPLGGGRHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMTYGT	449
Qy	481	LGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG	540
Db	450	LGGRLMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPG	509
Qy	541	VLLTRPVILAMDHCGEPSDPSWLSRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEAS	600
Db	510	VLLTRPVILAMDHCGEPSDPSWLSRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEAS	569
Qy	601	VFTEQLGRFALVGEALSVAALKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEV	660
Db	570	VFTEQLGRFALVGEALSVAALKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEV	629
Qy	661	KQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSCLLVSYQEIPFYHIWNGTQRY	720
Db	630	KQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSCLLVSYQEIPFYHIWNGTQRY	689
Qy	721	CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVP	780
Db	690	CTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVP	749
Qy	781	PSAFKIPFLIRQKIISLDPCCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWE	840
Db	750	PSAFKIPFLIRQKIISLDPCCRRGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWE	809
Qy	841	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	873
Db	810	RHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE	842

#### RESULT 4

##### UN5C\_MOUSE

ID UN5C\_MOUSE STANDARD; PRT; 931 AA.  
AC 008747; Q8CD16;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Netrin receptor UNC5C precursor (Unc-5 homolog C) (Unc-5 homolog 3)  
DE (Rostral cerebellar malformation protein).  
GN Name=Unc5c; Synonyms=Rcm, Unc5h3;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, DISEASE, AND TISSUE  
 RP SPECIFICITY.  
 RC STRAIN=C57B6/SJL;  
 RX MEDLINE=97271898; PubMed=9126743;  
 RA Ackerman S.L., Kozak L.P., Przyborski S.A., Rund L.A., Boyer B.B.,  
 RA Knowles B.B.;  
 RT "The mouse rostral cerebellar malformation gene encodes an UNC-5-like  
 RT protein.";  
 RL Nature 386:838-842(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusica V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [3]  
 RP FUNCTION, AND TISSUE SPECIFICITY.  
 RX PubMed=9389662;  
 RA Przyborski S.A., Knowles B.B., Ackerman S.L.;  
 RT "Embryonic phenotype of Unc5h3 mutant mice suggests chemorepulsion  
 RT during the formation of the rostral cerebellar boundary.";  
 RL Development 125:41-50(1998).  
 RN [4]  
 RP INTERACTION WITH DCC.  
 RX PubMed=10399920;  
 RA Hong K., Hinck L., Nishiyama M., Poo M.-M., Tessier-Lavigne M.,  
 RA Stein E.;

RT "A ligand-gated association between cytoplasmic domains of UNC5 and  
RT DCC family receptors converts netrin-induced growth cone attraction to  
RT repulsion.";  
RL Cell 97:927-941(1999).  
RN [5]  
RP PHOSPHORYLATION SITE TYR-568, AND MUTAGENESIS OF TYR-568.  
RX PubMed=11533026; DOI=10.1074/jbc.M103872200;  
RA Tong J., Killeen M., Steven R., Binns K.L., Culotti J., Pawson T.;  
RT "Netrin stimulates tyrosine phosphorylation of the UNC-5 family of  
RT netrin receptors and induces Shp2 binding to the RCM cytodomain.";  
RL J. Biol. Chem. 276:40917-40925(2001).  
RN [6]  
RP FUNCTION.  
RX PubMed=12451134; DOI=22/23/10346;  
RA Finger J.H., Bronson R.T., Harris B., Johnson K., Przyborski S.A.,  
RA Ackerman S.L.;  
RT "The netrin 1 receptors Unc5h3 and Dcc are necessary at multiple  
RT choice points for the guidance of corticospinal tract axons.";  
RL J. Neurosci. 22:10346-10356(2002).  
CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates  
CC axon repulsion of neuronal growth cones in the developing nervous  
CC system upon ligand binding. Axon repulsion in growth cones may be  
CC caused by its association with DCC that may trigger signaling for  
CC repulsion. Also involved in corticospinal tract axon guidances  
CC independently of DCC. It also acts as a dependence receptor  
CC required for apoptosis induction when not associated with netrin  
CC ligand.  
CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=008747-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=008747-2; Sequence=VSP\_011702;  
CC -!- TISSUE SPECIFICITY: Mainly expressed in regions of differentiating  
CC neurons. Highly expressed in brain and lung. Weakly expressed in  
CC testis, ovary, spleen, thymus and bladder. Expressed at very low  
CC level in kidney, intestine and salivary gland.  
CC -!- PTM: Phosphorylated on different cytoplasmic tyrosine residues.  
CC Phosphorylation of Tyr-568 leads to an interaction with PTPN11  
CC phosphatase, suggesting that its activity is regulated by  
CC phosphorylation/dephosphorylation. Tyrosine phosphorylation is  
CC netrin-dependent.  
CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The  
CC cleavage does not take place when the receptor is associated with  
CC netrin ligand. Its cleavage by caspases is required to induce  
CC apoptosis (By similarity).  
CC -!- DISEASE: Defects in Unc5c are the cause of rostral cerebellar  
CC malformation (Rcm). Rcm is characterized by cerebellar and  
CC midbrain defects, apparently as a result of abnormal neuronal  
CC migration.  
CC -!- SIMILARITY: Belongs to the UNC-5 family.  
CC -!- SIMILARITY: Contains 1 death domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.

```

CC      -!- SIMILARITY: Contains 1 ZU5 domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U72634; AAB54103.1; -.
DR      EMBL; AK031655; BAC27495.1; -.
DR      HSSP; P07996; 1LSL.
DR      MGD; MGI:1095412; Unc5c.
DR      GO; GO:0005886; C:plasma membrane; IC.
DR      GO; GO:0005042; F:netrin receptor activity; IDA.
DR      GO; GO:0005515; F:protein binding; IDA.
DR      GO; GO:0007420; P:brain development; IMP.
DR      GO; GO:0030334; P:regulation of cell migration; IMP.
DR      InterPro; IPR000488; Death.
DR      InterPro; IPR011029; DEATH_like.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR003598; Ig_c2.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      InterPro; IPR000906; ZU5.
DR      Pfam; PF00531; Death; 1.
DR      Pfam; PF00047; ig; 1.
DR      Pfam; PF00090; TSP_1; 2.
DR      Pfam; PF00791; ZU5; 1.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00005; DEATH; 1.
DR      SMART; SM00408; IGc2; 1.
DR      SMART; SM00209; TSP1; 2.
DR      SMART; SM00218; ZU5; 1.
DR      PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
DR      PROSITE; PS50835; IG_LIKE; 1.
DR      PROSITE; PS50092; TSP1; 2.
KW      Alternative splicing; Apoptosis; Developmental protein;
KW      Immunoglobulin domain; Phosphorylation; Receptor; Repeat; Signal;
KW      Transmembrane.
FT      SIGNAL          1      40      Potential.
FT      CHAIN           41     931     Netrin receptor UNC5C.
FT      DOMAIN          41     380     Extracellular (Potential).
FT      TRANSMEM        381     401     Potential.
FT      DOMAIN          402     931     Cytoplasmic (Potential).
FT      DOMAIN          62     159     Ig-like.
FT      DOMAIN          161     256     Ig-like C2-type.
FT      DOMAIN          260     314     TSP type-1 1.
FT      DOMAIN          316     368     TSP type-1 2.
FT      DOMAIN          528     631     ZU5.
FT      DOMAIN          850     929     Death.
FT      SITE            415     416     Cleavage (by caspase-3) (By similarity).
FT      SITE            694     712     Interaction with DCC (By similarity).
FT      DISULFID         83     142     By similarity.
FT      DISULFID        188     239     By similarity.
FT      MOD_RES         568     568     Phosphotyrosine.

```

FT	CARBOHYD	236	236	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	361	361	N-linked (GlcNAc. . .) (Potential).
FT	VARSP LIC	370	370	A -> GFIIYPISTEHRPQNEYGFSS (in isoform 2).
FT				/FTId=VSP_011702.
FT	MUTAGEN	568	568	Y->F: Abolishes interaction with PTPN11,
FT				leading to a increased level of
FT				phosphorylation.
FT	CONFLICT	16	16	L -> I (in Ref. 2).
FT	CONFLICT	733	733	H -> R (in Ref. 2).
FT	CONFLICT	924	924	S -> Y (in Ref. 2).
SQ	SEQUENCE	931 AA;	103062 MW;	8A5D951A4EECA179 CRC64;

Query Match 59.6%; Score 2778.5; DB 1; Length 931;  
 Best Local Similarity 58.2%; Pred. No. 2.2e-200;  
 Matches 513; Conservative 150; Mismatches 195; Indels 23; Gaps 7;

Qy	11	PGANPDLLPHFLVEPEDVYIVKKNPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG	70
Db	54	PSDPPEPLPHFLIEPEEAYIVKKNPVNLVCKASPATQIYFKCNSEWVHQKDHVVDERVD	113
Qy	71	SSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP	130
Db	114	TSGLIVREVSIEISRQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYLRKTFEQEP	173
Qy	131	LAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSPDNVYITREHSLVVRQARLAD	190
Db	174	LGKEVSLEQEVLLQCRPEGIPVAEVEWLKNEDIIDPAEDRNFYITIDHNLIIKQARLSD	233
Qy	191	TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPL	250
Db	234	TANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRTRTCTNPAPL	293
Qy	251	NGGAFCEGQNVQKTACATLCPVDGWSWSPWSKWSACGLDCTHWRSDPAPRNGGEECQ	310
Db	294	NGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTHWRRECTAPAPKNGGKDCD	353
Qy	311	GTDLDRNCTSDLCVHSASGPEDVALYVGL-IAVAVCLVLLLLLVILVYCRKKEGLDSDV	369
Db	354	GLVLQSKNCTDGLCMQAAPDSDDVALYVGIVIAVTVCLAITVVVAFVYRKNHRDFESDI	413
Qy	370	ADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSSTTTTYYQGSILCPQDGPSPKFQLTNGH	429
Db	414	IDSSALNGGFQPVNIKAARQD---LLAVPPDLTSAAAMYRGPVYALHD-VSDKIPMTNSP	469
Qy	430	LLSPLGGGRHTLHHSS----PTSEAEFVSRLS---TQNYF-----RSLPRGT---S	471
Db	470	ILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSQMTQSLLENEALNLKNQSLARQTDPS	529
Qy	472	NMTYGTFFNLGGRIMIPNTGISLLIPDAIPRGKIYEIYLT LHKPEDVRLPLAGCQTLLS	531
Db	530	CTAFGTFFNSLGGHLIPNSGVSLIPAGAIPOGRVYEMYVTVHRKENMRPPMEDSQTLLT	589
Qy	532	PIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCEGSWEDVLHLGEEAPSHLYY	591
Db	590	PVWSCGPPGALLTRPVILTLHHCADPSTEDWKIQKQAVQGWEDVVVGEENFTTPCY	649
Qy	592	CQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHD	651



CC -!- TISSUE SPECIFICITY: Mainly expressed in brain. Also expressed in  
CC kidney. Not expressed in developing or adult lung.  
CC -!- PTM: Phosphorylated on different cytoplasmic tyrosine residues:  
CC Phosphorylation of Tyr-568 leads to an interaction with PTPN11  
CC phosphatase, suggesting that its activity is regulated by  
CC phosphorylation/dephosphorylation. Tyrosine phosphorylation is  
CC netrin-dependent (By similarity).  
CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The  
CC cleavage does not take place when the receptor is associated with  
CC netrin ligand. Its cleavage by caspases is required to induce  
CC apoptosis.  
CC -!- DISEASE: Defects in Unc5c are the cause of cerebellar vermis  
CC defect (cvd) and hobble (hob) phenotypes. Cvd and hob rats exhibit  
CC cerebellar and midbrain defects, possibly as a result of abnormal  
CC neuronal migration, and exhibit laminar structure abnormalities in  
CC the fused cerebellar hemispheres and ectopic cerebellar tissues in  
CC the cerebello-pontine junction.  
CC -!- SIMILARITY: Belongs to the UNC-5 family.  
CC -!- SIMILARITY: Contains 1 death domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
CC -!- SIMILARITY: Contains 1 ZU5 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB118026; BAD05181.1; -.  
DR RGD; 735109; Unc5c.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR011029; DEATH\_like.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP\_1.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00531; Death; 1.  
DR Pfam; PF00047; ig; 1.  
DR Pfam; PF00090; TSP\_1; 2.  
DR Pfam; PF00791; ZU5; 1.  
DR PRINTS; PR01705; TSP1REPEAT.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00409; IG; 1.  
DR SMART; SM00408; IGc2; 1.  
DR SMART; SM00209; TSP1; 2.  
DR SMART; SM00218; ZU5; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; FALSE\_NEG.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
DR PROSITE; PS50092; TSP1; 2.  
KW Apoptosis; Developmental protein; Immunoglobulin domain;  
KW Phosphorylation; Receptor; Repeat; Signal; Transmembrane.



FT	SIGNAL	1	40	Potential.
FT	CHAIN	41	931	Netrin receptor UNC5C.
FT	DOMAIN	41	380	Extracellular (Potential).
FT	TRANSMEM	381	401	Potential.
FT	DOMAIN	402	931	Cytoplasmic (Potential).
FT	DOMAIN	62	159	Ig-like.
FT	DOMAIN	161	256	Ig-like C2-type.
FT	DOMAIN	260	314	TSP type-1 1.
FT	DOMAIN	316	368	TSP type-1 2.
FT	DOMAIN	528	631	ZU5.
FT	DOMAIN	850	929	Death.
FT	SITE	415	416	Cleavage (by caspase-3) (By similarity).
FT	SITE	694	712	Interaction with DCC (By similarity).
FT	DISULFID	83	142	By similarity.
FT	DISULFID	188	239	By similarity.
FT	MOD_RES	568	568	Phosphotyrosine (By similarity).
FT	CARBOHYD	236	236	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	361	361	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	931 AA;	103134 MW;	25B183A97BCB8401 CRC64;

Query Match 59.3%; Score 2762.5; DB 1; Length 931;  
 Best Local Similarity 57.9%; Pred. No. 3.5e-199;  
 Matches 510; Conservative 151; Mismatches 197; Indels 23; Gaps 7;

Qy	11	PGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG	70
Db	54	PSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQKDHVVDERVDE	113
Qy	71	SSGLPTMEVRINVSRRQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP	130
Db	114	TSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYLRKTTFEQEP	173
Qy	131	LAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLAD	190
Db	174	LGKEVSLEQEVLLQCRPPEGIPMAEVEWLKNEIDIIDPVEDRNFYITIDHNLIIKQARLSD	233
Qy	191	TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCCTNPAPL	250
Db	234	TANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWAEWSVCNSRCGRGYQKRTRTCTNPAPL	293
Qy	251	NGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWRSDPAPRNGGEECQ	310
Db	294	NGGAFCEGQSVQKIACTTLCPVDGRWTSWSKWSTCGTECTHWRRECTAPAPKNGGKDCD	353
Qy	311	GTDLDTNRCTSDLCVHSASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDV	369
Db	354	GLVLQSKNCTDGLCMQAAPSDDDVALYVGIVIAVTVCLAITVVVALFVYRKNHRDFESNI	413
Qy	370	ADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSSTTTTYQGSCLPRQDGPSPKFQLTNGH	429
Db	414	IDSSALNGGFQPVNIKAARQD---LLAVPPDLTSAAAMYRGPVYALHD-VSDKIPMTNSP	469
Qy	430	LLSPLGGGRHTLHHSS-----PTSEAEFVSRLS---TQNYF-----RSLPRGT--S	471
Db	470	ILDPLPNLKIKVYNSSGAVTPQDDLAEFSSKLSPPQMTQSLLENEALNLKNQSLARQTDPS	529
Qy	472	NMTYGTFFNFLGGRLMIPNTGISLLIPPDAIPRGKIYEIYLTLHKPEDVRLPLAGCQTLLS	531

```

      :|||| ||| |:|||:|:|||| | |||:|:|:|:|:|:|:|:| |:| | : ||||:
Db      530 CTAFGTFNSLGGHLIIPNSGVSLIIPAGAI PQGRVYEMYVTVHRKENMRPPMEDSQTLLT 589
Qy      532 PIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRLKKQSCGSWEDVLHLGEEAPSHLYY 591
      |:||||||| ||||||| : || :|: : | :|| |:| | ||||: :||| : |
Db      590 PVVSCGPPGALLTRPVILTLHHCADPNTEDWKIQ LKNQAVQGQWEDVVVVGEENFTTPCY 649
Qy      592 CQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHD 651
      ||:| ||:: || | :||||:: : ||||||| :| |: |:||||:||||| || |
Db      650 IQLDAEACHILTENLSTYALVGQSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQD 709
Qy      652 ALKEVVQLEKQLGGQLIQEPRLVHFKDSYHNLR LRSIHDVPSSLWKS KLLVSYQEIPFYHI 711
      ||||:||||:|:||||:|:|: |||| | |||||||: ||||||| |||||||
Db      710 ALKEVLQLERQMGGQLLEEPKALHFKGSIHNLR LRSIHDITHSLWKS KLLAKYQEIPFYHI 769
Qy      712 WNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALS 771
      |:|:| | |||||||:| :| :| || | ||||:| | :| ::: :| |:
Db      770 WSGSQRN LHCTFTLERLSINTVELVCKLCVRQVEGEGQIFQLNCTVSEPTGIDLPLLD 829
Qy      772 EAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLD SHLSFFASKPSPT 831
      : : : ||||| || ||||: |||| | || || | ||:| |:|:|:| |||
Db      830 ASTITTVTGPSAFSIP LPIRQKLCSSLDAPQTRGHDWRMLAHKLNLD RYLN YFATKSSPT 889
Qy      832 AMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 872
      :|:| ||||:|:|:| ||| : :|: : : :| :
Db      890 GVILDLWEAQNFDPGNLSMLAAVLEEMGRHETVVSLAAEGQ 930

```

# RESULT 6

## UN5C\_CHICK

```

ID      UN5C_CHICK      STANDARD;      PRT;      931 AA.
AC      Q7T2Z5;
DT      25-OCT-2004 (Rel. 45, Created)
DT      25-OCT-2004 (Rel. 45, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Netrin receptor UNC5C precursor (Unc-5 homolog C) (Unc-5 homolog 3)
DE      (cUNC-5H3).
GN      Name=UNC5C;
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
OX      NCBI_TaxID=9031;
RN      [1]
RP      SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX      PubMed=12799087;
RA      Guan W., Condic M.L.;
RT      "Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during
RT      chick dorsal root ganglia development.";
RL      Gene Expr. Patterns 3:369-373(2003).
CC      -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates
CC      axon repulsion of neuronal growth cones in the developing nervous
CC      system upon ligand binding (By similarity).
CC      -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC      -!- TISSUE SPECIFICITY: Restricted to proprioceptive neurons.
CC      -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
CC      similarity).

```

CC -!- SIMILARITY: Belongs to the UNC-5 family.  
CC -!- SIMILARITY: Contains 1 death domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
CC -!- SIMILARITY: Contains 1 ZU5 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AY187310; AAO67275.1; -.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00531; Death; 1.  
DR Pfam; PF00047; ig; 1.  
DR Pfam; PF00090; TSP\_1; 2.  
DR Pfam; PF00791; ZU5; 1.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00408; IGc2; 1.  
DR SMART; SM00209; TSP1; 2.  
DR SMART; SM00218; ZU5; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; FALSE\_NEG.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
DR PROSITE; PS50092; TSP1; 2.  
KW Developmental protein; Immunoglobulin domain; Phosphorylation;  
KW Receptor; Repeat; Signal; Transmembrane.  
FT SIGNAL 1 39 Potential.  
FT CHAIN 40 931 Netrin receptor UNC5C.  
FT DOMAIN 40 380 Extracellular (Potential).  
FT TRANSMEM 381 401 Potential.  
FT DOMAIN 402 931 Cytoplasmic (Potential).  
FT DOMAIN 62 159 Ig-like.  
FT DOMAIN 161 256 Ig-like C2-type.  
FT DOMAIN 260 314 TSP type-1 1.  
FT DOMAIN 316 368 TSP type-1 2.  
FT DOMAIN 528 631 ZU5.  
FT DOMAIN 850 929 Death.  
FT DISULFID 83 142 By similarity.  
FT DISULFID 188 239 By similarity.  
FT CARBOHYD 236 236 N-linked (GlcNAc . . .) (Potential).  
FT CARBOHYD 361 361 N-linked (GlcNAc . . .) (Potential).  
SQ SEQUENCE 931 AA; 102906 MW; 1E23A0D84F2E2C62 CRC64;

Query Match 59.0%; Score 2750.5; DB 1; Length 931;  
Best Local Similarity 57.8%; Pred. No. 2.8e-198;  
Matches 508; Conservative 150; Mismatches 198; Indels 23; Gaps 7;

Qy 11 PGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG 70  
| | : |||||:||||: ||||| | ||| |||||:|||| ||| | |||:: |

Db 54 PSDPPEPLPHFLIEPEEAYIVKNKPNVLYCKASPATQIYFKCNSEWVHQKDHVVDERVDE 113

Qy 71 SSGLPTEVRINVSRRQQVEKVFGLLEYYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP 130  
:||| || :|||||:| || :|||||:| |||:| |||:| ||| ||| |||

Db 114 TSGLIVCEVSIISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYLRKTFEQEP 173

Qy 131 LAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSPDPNVYITREHSLVVRQARLAD 190  
| ||||| :| ||||| |||||:|:| || | || :|:|:| |||:|

Db 174 LGKEVSLEQEVLLQCRPPEGIPVAEVEWLKNEEVIDPVEDRNFYITIDHNLIKQARLSD 233

Qy 191 TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPL 250  
|||||||:|:| :| ||||| ||||| ||:| |||:| |||:| |||||

Db 234 TANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSACNSRCGRGFQKRTRTCTNPAPL 293

Qy 251 NGGAFCEGQNVQKTACATLCPVDGWSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQ 310  
||||||| || ||||| |: |||| || :||| |||: |||:| |||:|

Db 294 NGGAFCEGQNVQKIACTTLCVPDGGKWTSSWSKSTCGTECTHWRRECTAPAPKNGGKDCE 353

Qy 311 GTDLDRNCTSDLCVHSASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDV 369  
| | :||| ||: | :|||||: || || : :| : || : ||:

Db 354 GLVLQSKNCTDGLCMQAAPSDDDVALYVGIVIAVIVCLAISVVVALFVYRKNHRDFESDI 413

Qy 370 ADSSILTSGFQPVSIKPSKADNPHLLTIQPDLTSTTTTQGSILCPQDGPSPKFQLTNGH 429  
||| | ||||:| || :| || : |||: | :| : | || :||

Db 414 IDSSALNGGFQPVNIKAARQD---LLAVPPDLTSAAAMYRGVPVYALHD-VSDKIPMTNSP 469

Qy 430 LLSPLGGRHTLHHSS----PTSEAEFVSRLS---TQNYF-----RSLPRGT--S 471  
:| || : :||| | | :| || ||: ||| :|| | |

Db 470 ILDPLPNLKIKVYNTSGAVTPQDELSDFSSKLSPOITQSLENETLNVKNQSLARQTDPS 529

Qy 472 NMFYGTFFNFLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLS 531  
:||| ||| :|||:| |||| | :|:|:|:|:|:| | :| | : |||:

Db 530 CTAFTFNSLGGHLVIPNSGVSLIPAGAVPQGRVYEMYVTVHRKEGMRPPVEDSQTLLT 589

Qy 532 PIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCGWSWEDVLHLGEEAPSHLYY 591  
|:||||| |||||:| | || ||: | | :|| | :| ||||: ||| : |

Db 590 PVVSCGPPGALLTRPVVLTMHHCAPNMDDWQIQLKHQAGQGPWEDVVVGEENFTTPCY 649

Qy 592 CQLEASACYVFTEQLGRFALVGEALSVAANKRLKLLLFAPVACTSLEYNIRVYCLHDTHD 651  
||: ||:| || | :||||:| ||||| | | :|:|:| ||||| || |

Db 650 IQLDPEACHILTETLSTYALVGQSITKAAAKRLKLAIFGPLSCSSLEYSIRVYCLDDTQD 709

Qy 652 ALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWKSLLVSYQEIPFYHI 711  
|||||:|:|:|:|:|:| ||| | |||||: ||||| |||||

Db 710 ALKEVLQLERQMGQLLEPKTLHFKGSTHNLRSLIHDIAHSLWKSCLPAKYQEIPFYHI 769

Qy 712 WNGTQRYLHCTFTLERVSPSTSDLACKLWWVQVEGDGQSFSINFNITKDTRFAELLALES 771  
|:| || ||||| | :| :| || | ||||:| | :| :||: : :|

Db 770 WSGCQRNLHCTFTLERFSLNTLELVCKLCVRQVEGEGQIFQLNCSVSEPTGIDYPIMDS 829

Qy 772 EAGVPALVGPSAFKIPFLIRQKISSLDPCCRRGADWRTLAQKLHLDShLSFFASKPSPT 831  
: :|||:| || ||||: |||| | || || | || || :|:|:| |||

Db 830 AGSITTIVGPNAFSIPLPIRQKLCSSLDAPQTRGHDWRMLAHKLKLDRLNYFATKSSPT 889

Qy 832 AMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSE 870  
:|:| |||:|:|:| ||| | : :| : : : |

Db 890 GVILDLWEAQNFDPGNLSMLAAVLEEMGRHETVVSIAAE 928

# RESULT 7

## UN5C\_HUMAN

ID UN5C\_HUMAN STANDARD; PRT; 931 AA.  
AC O95185; Q8IUT0;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Netrin receptor UNC5C precursor (Unc-5 homolog C) (Unc-5 homolog 3).  
GN Name=UNC5C; Synonyms=UNC5H3;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.  
RC TISSUE=Brain;  
RX MEDLINE=99000841; PubMed=9782087; DOI=10.1006/geno.1998.5425;  
RA Ackerman S.L., Knowles B.B.;  
RT "Cloning and mapping of the UNC5C gene to human chromosome 4q21-q23.";  
RL Genomics 52:205-208(1998).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Lung;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP DOWN-REGULATION IN CANCER.  
RX PubMed=12655055; DOI=10.1073/pnas.0738063100;  
RA Thiebault K., Mazelin L., Pays L., Llambi F., Joly M.-O.,  
RA Scoazec J.-Y., Saurin J.-C., Romeo G., Mehlen P.;  
RT "The netrin-1 receptors UNC5H are putative tumor suppressors  
RT controlling cell death commitment.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:4173-4178(2003).  
CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates  
CC axon repulsion of neuronal growth cones in the developing nervous  
CC system upon ligand binding. Axon repulsion in growth cones may be

CC caused by its association with DCC that may trigger signaling for  
 CC repulsion. Also involved in corticospinal tract axon guidances  
 CC independently of DCC. It also acts as a dependence receptor  
 CC required for apoptosis induction when not associated with netrin  
 CC ligand (By similarity).  
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=O95185-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=O95185-2; Sequence=VSP\_011700, VSP\_011701;  
 CC -!- TISSUE SPECIFICITY: Mainly expressed in brain. Also expressed in  
 CC kidney. Not expressed in developing or adult lung.  
 CC -!- PTM: Phosphorylated on different cytoplasmic tyrosine residues.  
 CC Phosphorylation of Tyr-568 leads to an interaction with PTPN11  
 CC phosphatase, suggesting that its activity is regulated by  
 CC phosphorylation/dephosphorylation. Tyrosine phosphorylation is  
 CC netrin-dependent (By similarity).  
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The  
 CC cleavage does not take place when the receptor is associated with  
 CC netrin ligand. Its cleavage by caspases is required to induce  
 CC apoptosis (By similarity).  
 CC -!- MISCELLANEOUS: Down-regulated in multiple cancers including  
 CC colorectal, breast, ovary, uterus, stomach, lung, or kidney  
 CC cancers.  
 CC -!- SIMILARITY: Belongs to the UNC-5 family.  
 CC -!- SIMILARITY: Contains 1 death domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC -!- SIMILARITY: Contains 1 ZU5 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF055634; AAC67491.1; -.  
 DR EMBL; BC041156; AAH41156.1; -.  
 DR HSSP; P07996; 1LSL.  
 DR Genew; HGNC:12569; UNC5C.  
 DR MIM; 603610; -.  
 DR GO; GO:0005042; F:netrin receptor activity; TAS.  
 DR GO; GO:0007411; P:axon guidance; TAS.  
 DR GO; GO:0007420; P:brain development; TAS.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR011029; DEATH\_like.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.

DR InterPro; IPR000906; ZU5.  
 DR Pfam; PF00531; Death; 1.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00090; TSP\_1; 2.  
 DR Pfam; PF00791; ZU5; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00408; IGc2; 1.  
 DR SMART; SM00209; TSP1; 2.  
 DR SMART; SM00218; ZU5; 1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; FALSE\_NEG.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS50092; TSP1; 2.  
 KW Alternative splicing; Apoptosis; Developmental protein;  
 KW Immunoglobulin domain; Phosphorylation; Polymorphism; Receptor;  
 KW Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 40 Potential.  
 FT CHAIN 41 931 Netrin receptor UNC5C.  
 FT DOMAIN 41 380 Extracellular (Potential).  
 FT TRANSMEM 381 401 Potential.  
 FT DOMAIN 402 931 Cytoplasmic (Potential).  
 FT DOMAIN 62 159 Ig-like.  
 FT DOMAIN 161 256 Ig-like C2-type.  
 FT DOMAIN 260 314 TSP type-1 1.  
 FT DOMAIN 316 368 TSP type-1 2.  
 FT DOMAIN 528 631 ZU5.  
 FT DOMAIN 850 929 Death.  
 FT SITE 415 416 Cleavage (by caspase-3) (By similarity).  
 FT SITE 694 712 Interaction with DCC (By similarity).  
 FT DISULFID 83 142 By similarity.  
 FT DISULFID 188 239 By similarity.  
 FT MOD\_RES 568 568 Phosphotyrosine (By similarity).  
 FT CARBOHYD 236 236 N-linked (GlcNAc. . .) (Potential).  
 FT CARBOHYD 361 361 N-linked (GlcNAc. . .) (Potential).  
 FT VARSPLIC 370 370 T -> SFIYPISTEQRTQNEYGFSS (in isoform 2).  
 FT /FTId=VSP\_011700.  
 FT VARSPLIC 579 931 Missing (in isoform 2).  
 FT /FTId=VSP\_011701.  
 FT VARIANT 37 37 G -> V (in dbSNP:2306715).  
 FT /FTId=VAR\_019731.  
 FT VARIANT 721 721 T -> M (in dbSNP:2289043).  
 FT /FTId=VAR\_019732.  
 FT CONFLICT 219 219 T -> I (in Ref. 1).  
 FT CONFLICT 489 489 S -> T (in Ref. 1).  
 SQ SEQUENCE 931 AA; 103101 MW; EFD71122C98DABB8 CRC64;

Query Match 59.0%; Score 2747.5; DB 1; Length 931;  
 Best Local Similarity 57.3%; Pred. No. 4.8e-198;  
 Matches 505; Conservative 151; Mismatches 202; Indels 23; Gaps 7;

Qy 11 PGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDG 70  
 | : |||||:||||: ||||| | ||| |||||:|||| ||| | ||::: |  
 Db 54 PSDPPEPLPHFLIEPEEAYIVKNKPVNLYCKASPATQIYFKCNSEWVHQKDHIVDERVDE 113  
 Qy 71 SSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP 130  
 :||| || | :|||||::|| |:|||||:|||||:||||:|||| ||| |||||  
 Db 114 TSGLIVREVSIEISRQQVEELFGPEDYWCQCVAWSSAGTTKSRKAYVRIAYLRKTFEQEP 173

Qy 131 LAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLAD 190  
| | | | | : | | | | | | | | | : | | | | : | | | : | | | | |  
Db 174 LGKEVSLEQEVLLQCRPEGIPVAEVEWLKNEDIIDPVEDRNFYITIDHNLIIKQARLSD 233  
Qy 191 TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCTNPAPL 250  
| | | | | : | | | | | | | | | : | | | | : | | | : | | | | |  
Db 234 TANYTCVAKNIVAKRKSTTATVIVYVNGGWSTWTEWSVCNSRCGRGYQKRTRTCTNPAPL 293  
Qy 251 NGGAFCEGQNVQKTACATLCPVDGWSWSPWSKWSACGLDCTHWRSSRECSDPAPRNGGEECQ 310  
| | | | | : | | | | | | | | | : | | | | : | | | : | | | : |  
Db 294 NGGAFCEGQSVQKIACTTLCVPDGRWTPWSKWSTCGTECTHWRRRRECTAPAPKNGGKDCD 353  
Qy 311 GTDLDTNRCTSDLCVHSASGPEDVALYVGL-IAVAVCLVLLLLVLILVYCRKKEGLDSDV 369  
| | : | | | | : | | : | | | | : | | | : | | : | | : | | :  
Db 354 GLVLQSKNCTDGLCMQTAPDSDVALYVGIVIAVIVCLAISVVVALFVYRKNHRDFESDI 413  
Qy 370 ADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSSTTTTTYQGSICPRQDGPSPKFQLTNGH 429  
| | | | | : | | : | | | | : | | : | | : | | : | | : | |  
Db 414 IDSSALNGGFQPVNIKAARQD---LLAVPPDLTSAAAMYRGPVYALHD-VSDKIPMTNSP 469  
Qy 430 LLSPLGGGRHTLHHS----SPTSEAEFVSRLS---TQNYF-----RSLPRGT--S 471  
: | | : : : | | : | | : | | : | | : | | : | | : | |  
Db 470 ILDPLPNLKIKVYNTSGAVSPQDDLSEFTSKLSPQMTQSLLENEALSLKNQSLARQTDPS 529  
Qy 472 NMTYGTFFNFLGGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLAGCQTLLS 531  
: | | | | : | | : | | : | | : | | : | | : | | : | | : | |  
Db 530 CTAFGSFNSLGGHLIVPNSGVSLIPAGAI PQGRVYEMYVTVHRKETMRPPMDDSQTLT 589  
Qy 532 PIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLRLKKQSCEGSWEDVLHLGEEAPSHLYY 591  
| : | | | | | | | | | : | | : | : | : | | : | | : | | : | |  
Db 590 PVVSCGPPGALLTRPVVLTMHHCADPNTEDWKILLKNQAAQGWEDVVVVGEENFTTPCY 649  
Qy 592 CQLEASACYVFTEQLGRFALVGEALSVA AAKRLKLLLFAPVACTSLEYNIRVYCLHDTHD 651  
: | : | | : : | | : | | : : | | | | | : | | : | | : | | | | |  
Db 650 IKLDAEACHILTENLSTYALVGHSTTKAAAKRLKLAIFGPLCCSSLEYSIRVYCLDDTQD 709  
Qy 652 ALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLRSLIHDVPSSLWWSKLLVSYQEIPFYHI 711  
| | | : : | | : | | : : | | : | | : | | : | | : | | : | | :  
Db 710 ALKEILHLERQTGGQLLEPKALHFKGSTHNLRSLIHDIAHSLWWSKLLAKYQEIPFYHV 769  
Qy 712 WNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALES 771  
| : | : | | | | | | | : | : | | | | | | | : | : : : : | | :  
Db 770 WSGSQRNLHCTFTLERFSLNTVELVCKLCVRQVEGEGQIFQLNCTVSEEPTGIDLPLLD 829  
Qy 772 EAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQKLHLDShLSFFASKPSPT 831  
: : | | | | | | | : | | : | | | | | | | : | | : : : | | |  
Db 830 ANTITTVTGPSAFSIPPIRQKLCSSLDAPQTRGHDWRMLAHKLNLDRLNYFATKSSPT 889  
Qy 832 AMILNLWEARHFPNGNLSQLAAAVAGLGQPDAGLFTVSEAE 872  
: | : | | : : | | | | | : : : : : : : : : : : : : : :  
Db 890 GVILDLWEAQNFDPGNLSMLAAVLEEMGRHETVVSLAAEGQ 930

RESULT 8

UN5B\_XENLA

ID UN5B\_XENLA

STANDARD;

PRT;

943 AA.



AC Q8JGT4;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Netrin receptor UNC5B precursor (UNC-5 homolog) (Protein XUNC-5).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RX MEDLINE=22239703; PubMed=12351179; DOI=10.1016/S0925-4773(02)00215-0;  
RA Anderson R.B., Holt C.E.;  
RT "Expression of UNC-5 in the developing Xenopus visual system.";  
RL Mech. Dev. 118:157-160(2002).  
CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates  
CC axon repulsion of neuronal growth cones in the developing nervous  
CC system upon ligand binding (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- TISSUE SPECIFICITY: In the developing visual system, it is  
CC expressed within the developing optic vesicles and later become  
CC restricted to the dorsal ciliary marginal zone, a site of  
CC retinoblast proliferation and differentiation.  
CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By  
CC similarity).  
CC -!- SIMILARITY: Belongs to the UNC-5 family.  
CC -!- SIMILARITY: Contains 1 death domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
CC -!- SIMILARITY: Contains 1 ZU5 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AY099459; AAM34486.1; -.  
DR HSSP; P07996; 1LSL.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR011029; DEATH\_like.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP\_1.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00531; Death; 1.  
DR Pfam; PF00047; ig; 1.  
DR Pfam; PF00090; TSP\_1; 2.  
DR Pfam; PF00791; ZU5; 1.  
DR PRINTS; PR01705; TSP1REPEAT.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00408; IGc2; 1.



```

      |:|::  || : | :|| || || : ::||
Db      453 GIYRGNMYALQDS-ADKIPMTNSPLLDPLPNLKIKVYNSSTVGSSPGIHDGNLLGTKPT 511.
Qy      448 SEAEFVSRLSTQN-----YFRSLPRGTSNMTYGTFFNLGGRIMIPNTGISLLIPDAI 501
      : :: :| : :|| :|| ||| |||| |||:|||| ||
Db      512 GTYPSDNNIMNARNKNMSMQHLLTLPRDSSNSVTGTFGSLGGRLTFPNTGVSLIPQGA 571
Qy      502 PRGKIYEIYLT LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPSDS 561
      |:|| ||:|| ::| |: || | ||:||||:|||| |:|| :|||| : || : :
Db      572 PQGKYEMYLMINKRENTVLPSEGTQTILSPIITCGPTGLLLCCKPVILTVPHCADINTSD 631
Qy      562 WSLRLKKQSCGWSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 621
      | |:|| || :|:|:|:| | || : | ||||: :: : ||| :| |||: | :|
Db      632 WILQLKTQSHQGNWEEVVTLINEETLNTPCYQCLESCHTLLDQLGTYAFVGESYSRSAI 691
Qy      622 KRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYH 681
      |||:| :||: | |||||:|:|: || | ||||:|:| || | |:|:|:| |||||
Db      692 KRLQLAIFAPMLCTSLEYNLKVYCVEDTPDALKEVLELEKTLGGYLVEEPKLLMFKDSYH 751
Qy      682 NLRLSIHDVPSSLWWSKLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWV 741
      |||||:| | |:|:|:| | |||||:|:|: || | |||||:|:| || | |:|:|:| ||| |
Db      752 NLRLSIHDIPHSLSWRSKLMAKYQEIPFYHIWSGSQRTLHCTFTLERYSLAATELTCKICV 811
Qy      742 WQVEGDGQSFSINFNITKDRFAELLALESEAGVPALVGPSAFKIPFLIRQKIISLDP 801
      |||:| | | : : : : : : :| | :| ||||| |||| :|| |
Db      812 RQVEGEGQIFQLHTLLEENVKSFDPFCSQAENSVTTHLGPYAFKIPFSIRQKICNSLDAP 871
Qy      802 CRRGADWRTLAQKLHLDLSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQP 861
      || ||| ||||| :| :|:|:|:| ||| :|:|:| || | :|:|: ||:|: :|:
Db      872 NSRGNDWRLLAQKLCMDRYLNYFATKASPTGVILDLWEALHQDDGDLNTIASALEEMGKS 931
Qy      862 DAGLFTVSEAE 873
      : | : : :|
Db      932 EMLVMATDGDC 943

```

# RESULT 9

## UN5B\_MOUSE

```

ID   UN5B_MOUSE          STANDARD;          PRT;          945 AA.
AC   Q8K1S3; Q6PFH0; Q80Y85; Q9D398;
DT   25-OCT-2004 (Rel. 45, Created)
DT   25-OCT-2004 (Rel. 45, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Netrin receptor UNC5B precursor (Unc-5 homolog B) (Unc-5 homolog 2).
GN   Name=Unc5b; Synonyms=Unc5h2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RX   MEDLINE=22239710; PubMed=12351186; DOI=10.1016/S0925-4773(02)00248-4;
RA   Engelkamp D.;
RT   "Cloning of three mouse unc-5 genes and their expression patterns at
RT   mid-gestation.";
RL   Mech. Dev. 118:191-197(2002).
RN   [2]

```

RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusica V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasaki Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempole C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).

RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]

RP TISSUE SPECIFICITY.  
RX PubMed=12799072;  
RA Dalvin S., Anselmo M.A., Prodhan P., Komatsuzaki K., Schnitzer J.J.,  
RA Kinane T.B.;  
RT "Expression of Netrin-1 and its two receptors DCC and UNC5H2 in the  
RT developing mouse lung.";  
RL Gene Expr. Patterns 3:279-283(2003).  
CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates  
CC axon repulsion of neuronal growth cones in the developing nervous  
CC system upon ligand binding. Axon repulsion in growth cones may be  
CC caused by its association with DCC that may trigger signaling for  
CC repulsion. It also acts as a dependence receptor required for  
CC apoptosis induction when not associated with netrin ligand (By  
CC similarity).  
CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC. Interacts  
CC with GNAI2 via its cytoplasmic part (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=Q8K1S3-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=Q8K1S3-2; Sequence=VSP\_011699;  
CC -!- TISSUE SPECIFICITY: Highly expressed in brain. Expressed in lung  
CC during late development. Expressed during early blood vessel  
CC formation, in the semicircular canal and in a dorsal to ventral  
CC gradient in the retina.  
CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By  
CC similarity).  
CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The  
CC cleavage does not take place when the receptor is associated with  
CC netrin ligand. Its cleavage by caspases is required to induce  
CC apoptosis (By similarity).  
CC -!- SIMILARITY: Belongs to the UNC-5 family.  
CC -!- SIMILARITY: Contains 1 death domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
CC -!- SIMILARITY: Contains 1 ZU5 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AJ487853; CAD32251.1; -.  
DR EMBL; AK018177; BAB31108.1; -.  
DR EMBL; BC048162; AAH48162.1; ALT\_INIT.  
DR EMBL; BC057560; AAH57560.1; -.  
DR HSSP; P07996; 1LSL.  
DR MGD; MGI:894703; Unc5b.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR011029; DEATH\_like.  
DR InterPro; IPR007110; Ig-like.

```

DR      InterPro; IPR003598; Ig_c2.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      InterPro; IPR000906; ZU5.
DR      Pfam; PF00531; Death; 1.
DR      Pfam; PF00047; ig; 1.
DR      Pfam; PF00090; TSP_1; 2.
DR      Pfam; PF00791; ZU5; 1.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00005; DEATH; 1.
DR      SMART; SM00408; IGc2; 1.
DR      SMART; SM00209; TSP1; 2.
DR      SMART; SM00218; ZU5; 1.
DR      PROSITE; PS50017; DEATH_DOMAIN; 1.
DR      PROSITE; PS50835; IG_LIKE; 1.
DR      PROSITE; PS50092; TSP1; 2.
KW      Alternative splicing; Apoptosis; Developmental protein;
KW      Immunoglobulin domain; Phosphorylation; Receptor; Repeat; Signal;
KW      Transmembrane.
FT      SIGNAL      1      26      Potential.
FT      CHAIN      27      945      Netrin receptor UNC5B.
FT      DOMAIN      27      377      Extracellular (Potential).
FT      TRANSMEM      378      398      Potential.
FT      DOMAIN      399      945      Cytoplasmic (Potential).
FT      DOMAIN      48      145      Ig-like.
FT      DOMAIN      153      242      Ig-like C2-type.
FT      DOMAIN      246      300      TSP type-1 1.
FT      DOMAIN      302      354      TSP type-1 2.
FT      DOMAIN      541      644      ZU5.
FT      DOMAIN      865      943      Death.
FT      SITE      412      413      Cleavage (by caspase-3) (By similarity).
FT      SITE      707      725      Interaction with DCC (By similarity).
FT      DISULFID      69      128      By similarity.
FT      DISULFID      174      225      By similarity.
FT      CARBOHYD      222      222      N-linked (GlcNAc. . .) (Potential).
FT      CARBOHYD      347      347      N-linked (GlcNAc. . .) (Potential).
FT      VARSPLIC      356      367      NQRTLNDPKSHSP -> T (in isoform 2).
FT      /FTId=VSP_011699.
FT      CONFLICT      238      238      T -> A (in Ref. 2).
FT      CONFLICT      394      394      V -> E (in Ref. 2).
FT      CONFLICT      679      679      T -> S (in Ref. 2).
FT      CONFLICT      874      874      N -> D (in Ref. 2).
SQ      SEQUENCE      945 AA: 103738 MW: 80E896F0F0E06012 CRC64:

```

Db 160 LAKEVPLDHEVLLQCRPPEGVPVAEVEWLKNEVDIDPAQDTNFLTIDHNLIIIRQARLSD 219  
 Qy 191 TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSTNPAPL 250  
 Db 220 TANYTCVAKNIVAKRRSTTATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRTCTNPAPL 279  
 Qy 251 NGGAFCEGQNVQKTACATLCPVDGSWSPWSKWSACGLDCTHWSRECSDPAPRNGGEECQ 310  
 Db 280 NGGAFCEGQAFQKTACTTVCPVDGAWTEWSKWSACSTECAHWSRECMAPPPQNGGRDCS 339  
 Qy 311 GTDLDTNRCTSDLCV-----HSASGPEDVALYVGL-IAVAVCLVLLLLVLILVY 358  
 Db 340 GTLLDSKNCTDGLCVLNQRTLNDPKSHPLETSGDVALYAGLVVAVFVVAVLMAVGIVY 399  
 Qy 359 CRKKEGLDSDVADSS-ILTSGFQPVSIKPSKADNPHLL--TIQPDLSSTTTTYQGS LCPR 415  
 Db 400 RRNCRDFDITDSSAALTGGFHPVNFKTARPNNPQLLHPSAPPDLTASAGIYRGPVYAL 459  
 Qy 416 QDGPSPKFQLTNGHLLSPLGGGRHTLHHSS-----PTSEAEFEVFSR 456  
 Db 460 QDS-ADKIPMTNSPLLDPLPSLKIKVYNSSTIGSGSGLADGADLLGVLPPGTYPGDF-SR 517  
 Qy 457 LSTQNYFRS-----LPRGTSNMTYGTFFNLGGRLMIPNTGISLLIPPDAIPRGKI 506  
 Db 518 DTHFLHRSASLGSQHLLGLPRDPSSSVSGTFGCLGGRSLPGTGVSLIVPENGAIPOGKF 577  
 Qy 507 YEIYLTLHKPEDVRLPLA-GCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWSLR 565  
 Db 578 YDLYLHINKAEST-LPLSEGSQTVLSPSVTCGPTGLLLCPVVLTVPHCAEVIAGDWIFQ 636  
 Qy 566 LKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA AAKRLK 625  
 Db 637 LKTQAHQGHWEVVTLD EETLNTPCYCQLEAKSCHILLDQLGTYVFMGESYSRSAVKRLQ 696  
 Qy 626 LLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNLR 685  
 Db 697 LAIFAPALCTSLEYSRLVYCLEDT PVALKEVLELERTLGGYLVEEPKPLLFKDSYHNLR 756  
 Qy 686 SIHDVPSSLWKS KLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSD LACKLWWQVE 745  
 Db 757 SLHDIPHAHWSKLLAKYQEIPFYHVWNGSQRALHCTFTLERHSLASTEFTCKVCVRQVE 816  
 Qy 746 GDGQSFSINFNITKDTRFAELLALESEAG--VPALVGPSAFKIPFLIRQKIISSLDPPCR 803  
 Db 817 GEGQIFQLHTTLA-ETPAGSLDALCSAPGNAIT TQLGPYAFKIPLSIRQKICSSLDAPNS 875  
 Qy 804 RGADWRTLAQKLHLD SHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPD 863  
 Db 876 RGNDWRLLAQKLSMDRYLNYFATKASPTGVILD LWEARQQDDGDLNSLASALEEMGKSEM 935  
 Qy 864 GLFTVSEAE 873  
 Db 936 LVAMATDGDC 945

## UN5B\_RAT

ID UN5B\_RAT STANDARD; PRT; 945 AA.  
AC 008722;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Netrin receptor UNC5B precursor (Unc-5 homolog B) (Unc-5 homolog 2).  
GN Name=Unc5b; Synonyms=Unc5h2;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE  
RP SPECIFICITY.  
RX MEDLINE=97271897; PubMed=9126742;  
RA Leonardo E.D., Hinck L., Masu M., Keino-Masu K., Ackerman S.L.,  
RA Tessier-Lavigne M.;  
RT "Vertebrate homologues of C. elegans UNC-5 are candidate netrin  
RT receptors.";  
RL Nature 386:833-838(1997).  
RN [2]  
RP FUNCTION, AND INTERACTION WITH DCC.  
RX PubMed=10399920;  
RA Hong K., Hinck L., Nishiyama M., Poo M.-M., Tessier-Lavigne M.,  
RA Stein E.;  
RT "A ligand-gated association between cytoplasmic domains of UNC5 and  
RT DCC family receptors converts netrin-induced growth cone attraction to  
RT repulsion.";  
RL Cell 97:927-941(1999).  
RN [3]  
RP FUNCTION, AND MUTAGENESIS OF ASP-412.  
RX PubMed=11387206; DOI=10.1093/emboj/20.11.2715;  
RA Llambi F., Causeret F., Bloch-Gallego E., Mehlen P.;  
RT "Netrin-1 acts as a survival factor via its receptors UNC5H and DCC.";  
RL EMBO J. 20:2715-2722(2001).  
CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates  
CC axon repulsion of neuronal growth cones in the developing nervous  
CC system upon ligand binding. Axon repulsion in growth cones may be  
CC caused by its association with DCC that may trigger signaling for  
CC repulsion. It also acts as a dependence receptor required for  
CC apoptosis induction when not associated with netrin ligand.  
CC -!- SUBUNIT: Interacts with GNAI2 via its cytoplasmic part (By  
CC similarity). Interacts with the cytoplasmic part of DCC.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Mainly expressed in regions of differentiating  
CC neurons. Expressed in the developing sensory ganglia that flank  
CC the spinal cord from E12, peaking at E14. Expressed in the roof  
CC plate region of the spinal cord from E14.  
CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By  
CC similarity).  
CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The  
CC cleavage does not take place when the receptor is associated with  
CC netrin ligand. Its cleavage by caspases is required to induce  
CC apoptosis.  
CC -!- SIMILARITY: Belongs to the UNC-5 family.  
CC -!- SIMILARITY: Contains 1 death domain.



CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
CC -!- SIMILARITY: Contains 1 ZU5 domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U87306; AAB57679.1; -.  
DR HSSP; P07996; 1LSL.  
DR RGD; 621756; Unc5h2.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR011029; DEATH\_like.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP\_1.  
DR InterPro; IPR000906; ZU5.  
DR Pfam; PF00531; Death; 1.  
DR Pfam; PF00047; ig; 1.  
DR Pfam; PF00090; TSP\_1; 2.  
DR Pfam; PF00791; ZU5; 1.  
DR PRINTS; PR01705; TSP1REPEAT.  
DR SMART; SM00005; DEATH; 1.  
DR SMART; SM00408; IGc2; 1.  
DR SMART; SM00209; TSP1; 2.  
DR SMART; SM00218; ZU5; 1.  
DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
DR PROSITE; PS50835; IG\_LIKE; 1.  
DR PROSITE; PS50092; TSP1; 2.  
KW Apoptosis; Developmental protein; Immunoglobulin domain;  
KW Phosphorylation; Receptor; Repeat; Signal; Transmembrane.  
FT SIGNAL 1 26 Potential.  
FT CHAIN 27 945 Netrin receptor UNC5B.  
FT DOMAIN 27 377 Extracellular (Potential).  
FT TRANSMEM 378 398 Potential.  
FT DOMAIN 399 945 Cytoplasmic (Potential).  
FT DOMAIN 48 145 Ig-like.  
FT DOMAIN 153 242 Ig-like C2-type.  
FT DOMAIN 246 300 TSP type-1 1.  
FT DOMAIN 302 354 TSP type-1 2.  
FT DOMAIN 541 644 ZU5.  
FT DOMAIN 865 943 Death.  
FT SITE 412 413 Cleavage (by caspase-3).  
FT SITE 707 725 Interaction with DCC.  
FT DISULFID 69 128 By similarity.  
FT DISULFID 174 225 By similarity.  
FT CARBOHYD 222 222 N-linked (GlcNAc . . .) (Potential).  
FT CARBOHYD 347 347 N-linked (GlcNAc . . .) (Potential).  
FT MUTAGEN 412 412 D->N: Abolishes cleavage by caspase-3 and  
FT subsequent induction of apoptosis.  
SQ SEQUENCE 945 AA; 103520 MW; 6E9C2A262E560B9B CRC64;

Query Match 54.9%; Score 2560.5; DB 1; Length 945;  
Best Local Similarity 54.0%; Pred. No. 6.2e-184;  
Matches 496; Conservative 139; Mismatches 215; Indels 69; Gaps 14;

Qy	11	PGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKNGEWMVRQVDHVIERSTDG	70
Db	40	PSAPAEQLPHFLLEPEDAYIVKNKPVLELHCRAFPATQIYFKNGEWMVVSQKGHVTDQESLDE	99
Qy	71	SSGLPTMEVRINVSRRQQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP	130
Db	100	ATGLRIREVQIEVSRQQVEELFGLLEDYWCQCVAWSSSGTTKSRRAYIRIAYLRKNFDQEP	159
Qy	131	LAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDPSPDPNVYITREHSLVVRQARLAD	190
Db	160	LAKEVPLDHEVLLQCRPPGVPVAEVEWLKNEVDIDPAQDTNFLTIDHNLIIQRARLSD	219
Qy	191	TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAPL	250
Db	220	TANYTCVAKNIVAKRRSTTATVIVYVNGGWSSWAEWSPCSNRCGRGWQKRTRCTNTPAPL	279
Qy	251	NGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWRSDPAPRNGGEECQ	310
Db	280	NGGAFCEGQACQKTACTTVCPVDGAWTEWSKWSACSTECAHWRSDRECMAPPQNGGRDCS	339
Qy	311	GTDLDRNCTSDLCV---HSASGPE-----DVALYVGL-IAVAVCLVLLLLVLILVY	358
Db	340	GTLLDSKNCTDGLCVLNQRTLNDPKSRPLEPSGDVALYAGLVVAVFVVLAVLMAVGIVY	399
Qy	359	CRKKEGLDSDVADSS-ILTSGFQPVSIKPSKADNPHLL--TIQPDLSSTTTTTYQGSCLPR	415
Db	400	RRNCRDFDITDITDSSAALTGGFHPVNFKTARPSNPQLLHPSAPPDLTASAGIYRGPVYAL	459
Qy	416	QDGSPKPFQLTNHLLSPL-----GGG-----	437
Db	460	QDS-ADKIPMTNSPLLDPLPSLKIKVYDSSSTIGSGAGLADGADLLGVLPPGTYPGDFSRD	518
Qy	438	RHTLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMITYGTNFLGGRLMIPNTGISLLIP	497
Db	519	THFLHLRS-----ASLGSQ-HLLGLPRDPSSSVSGTFGCLGGRLTIPGTGVSLLV	568
Qy	498	PDAIPRGKIYEIYLTLHKPEDVRLPLA-GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGE	556
Db	569	NGAIPQGFYDLYLRINKTEST-LPLSEGSQTVLSPSVTCGPTGLLLCRPVVLTVPHCAE	627
Qy	557	PSPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL	616
Db	628	VIAGDWIFQLKTAHQHGWEEVVTLDDETLNTPCYCQLEAKSCHILLDQLGTYVFTGESY	687
Qy	617	SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF	676
Db	688	SRSVAKRLQLAIFAPALCTSLEYSLRVYCLEDTPAALKEVLELERTLGGYLVEEPKTLF	747
Qy	677	KDSYHNLRSLIHDPVSSLWKSLLVSYQEIIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA	736
Db	748	KDSYHNLRSLHDIPHAHWRSKLLAKYQEIIPFYHVWNGSOKALHCTFTLERHSLASTEFT	807

Qy 737 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPAL--VGPSAFKIPFLIRQKI 794  
 ||: | |||: || | : : : | | | | | | : || |||| | ||||  
 Db 808 CKVCVRQVEGEGQIFQLHTTLA-ETPAGSLDALCSAPGNAATTQLGPYAFKIPLSIRQKI 866  
 Qy 795 ISSLDPPCRRGADWRTLQAQKLHLSHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAA 854  
 : || | | | || | |||| : | : ||: ||: | || | : ||: |||| : ||: ||: |  
 Db 867 CNSLDAPNSRGNDWRLLAQKLSMDRYLNYFATKASPTGVILDLWEARQQDDGDLNSLASA 926  
 Qy 855 VAGLGQPDAGLFTVSEAEC 873  
 : : ||: : : : : : |  
 Db 927 LEEMGKSEMLVAMTTDGDGDC 945

# RESULT 11

## UN5B\_HUMAN

ID UN5B\_HUMAN STANDARD; PRT; 945 AA.  
 AC Q8IZJ1; Q86SN3; Q8N1Y2; Q9H9F3;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Netrin receptor UNC5B precursor (Unc-5 homolog B) (Unc-5 homolog 2)  
 DE (p53-regulated receptor for death and life protein 1)  
 DE (UNQ1883/PRO4326).  
 GN Name=UNC5B; Synonyms=P53RDL1, UNC5H2;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), TISSUE SPECIFICITY, AND INTERACTION  
 RP WITH GNAI2.  
 RC TISSUE=Lung;  
 RX MEDLINE=22246081; PubMed=12359238; DOI=10.1016/S0006-291X(02)02277-5;  
 RA Komatsuzaki K., Dalvin S., Kinane T.B.;  
 RT "Modulation of G(ialpha(2)) signaling by the axonal guidance molecule  
 RT UNC5H2.";  
 RL Biochem. Biophys. Res. Commun. 297:898-905(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), FUNCTION, AND MUTAGENESIS OF ASP-412.  
 RX PubMed=12598906; DOI=10.1038/ncb943;  
 RA Tanikawa C., Matsuda K., Fukuda S., Nakamura Y., Arakawa H.;  
 RT "p53RDL1 regulates of p53-dependent apoptosis.";  
 RL Nat. Cell Biol. 5:216-223(2003).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P., Gray A.;  
 RT "The secreted protein discovery initiative (SPDI), a large-scale  
 RT effort to identify novel human secreted and transmembrane proteins: a

RT bioinformatics assessment.";  
 RL Genome Res. 13:2265-2270(2003).  
 RN [4]  
 RP SEQUENCE OF 361-945 FROM N.A.  
 RC TISSUE=Amygdala, and Teratocarcinoma;  
 RX PubMed=14702039; DOI=10.1038/ngl285;  
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,  
 RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,  
 RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,  
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,  
 RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,  
 RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,  
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,  
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,  
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,  
 RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,  
 RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,  
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,  
 RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,  
 RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,  
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,  
 RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,  
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,  
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,  
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,  
 RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,  
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,  
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,  
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,  
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,  
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,  
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,  
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;  
 RT "Complete sequencing and characterization of 21,243 full-length human  
 RT cDNAs.";  
 RL Nat. Genet. 36:40-45(2004).  
 RN [5]  
 RP DOWN-REGULATION IN CANCER.  
 RX PubMed=12655055; DOI=10.1073/pnas.0738063100;  
 RA Thiebault K., Mazelin L., Pays L., Llambi F., Joly M.-O.,  
 RA Scoazec J.-Y., Saurin J.-C., Romeo G., Mehlen P.;  
 RT "The netrin-1 receptors UNC5H are putative tumor suppressors  
 RT controlling cell death commitment.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:4173-4178(2003).  
 CC -!- FUNCTION: Receptor for netrin required for axon guidance. Mediates  
 CC axon repulsion of neuronal growth cones in the developing nervous  
 CC system upon ligand binding. Axon repulsion in growth cones may be  
 CC caused by its association with DCC that may trigger signaling for  
 CC repulsion. It also acts as a dependence receptor required for  
 CC apoptosis induction when not associated with netrin ligand.  
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC (By  
 CC similarity). Interacts with GNAI2 via its cytoplasmic part.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;

```

CC      IsoId=Q8IZJ1-1; Sequence=Displayed;
CC      Name=2;
CC      IsoId=Q8IZJ1-2; Sequence=VSP_011698;
CC      -!- TISSUE SPECIFICITY: Highly expressed in brain. Also expressed at
CC          lower level in developing lung, cartilage, kidney and
CC          hematopoietic and immune tissues.
CC      -!- INDUCTION: By p53/TP53.
CC      -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
CC          similarity).
CC      -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
CC          cleavage does not take place when the receptor is associated with
CC          netrin ligand. Its cleavage by caspases is required to induce
CC          apoptosis.
CC      -!- MISCELLANEOUS: Down-regulated in multiple cancers including
CC          colorectal, breast, ovary, uterus, stomach, lung, or kidney
CC          cancers.
CC      -!- SIMILARITY: Belongs to the UNC-5 family.
CC      -!- SIMILARITY: Contains 1 death domain.
CC      -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC      -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC      -!- SIMILARITY: Contains 2 TSP type-1 domains.
CC      -!- SIMILARITY: Contains 1 ZU5 domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AY126437; AAM95701.1; -.
DR      EMBL; AB096256; BAC57998.1; -.
DR      EMBL; AY358351; AAQ88717.1; -.
DR      EMBL; AK022859; BAB14276.1; ALT_INIT.
DR      EMBL; AK094595; BAC04382.1; ALT_INIT.
DR      HSSP; P07996; 1LSL.
DR      Genew; HGNC:12568; UNC5B.
DR      MIM; 607870; -.
DR      InterPro; IPR000488; Death.
DR      InterPro; IPR011029; DEATH_like.
DR      InterPro; IPR007110; Ig-like.
DR      InterPro; IPR003598; Ig_c2.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      InterPro; IPR000906; ZU5.
DR      Pfam; PF00531; Death; 1.
DR      Pfam; PF00047; ig; 1.
DR      Pfam; PF00090; TSP_1; 2.
DR      Pfam; PF00791; ZU5; 1.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00005; DEATH; 1.
DR      SMART; SM00408; IGc2; 1.
DR      SMART; SM00209; TSP1; 2.
DR      SMART; SM00218; ZU5; 1.
DR      PROSITE; PS50017; DEATH_DOMAIN; FALSE_NEG.
DR      PROSITE; PS50835; IG_LIKE; 1.

```



Db 332 QNGGRDCSGTLLDSKNCTDGLCMQNKKTLSDPNSHLLEASG--DAALYAGLVVAIFVWVA 389  
 Qy 349 LLLLVLILVYCRKKEGLSDVADSS-ILTSGFQPVSIKPSKADNPHELL--TIQPDLSSTTT 405  
 Db 390 ILMAVGVVYRRNCRDFDITDSSAALTGGFHPVNFKTARPSNPQLLHPSVPPDLTASA 449  
 Qy 406 TTYQGSGLCPQDGPSPKFQLTNGHLLSPLGGGRHTLHHSSPT-----SEAEFVSRLST 459  
 Db 450 GIYRGPVYALQDS-TDKIPMTNSPLLDPLPSLKVKVYSSSTTGSGPGLADGADLLGLVLP 508  
 Qy 460 QNY-----FRS-----LPRGTSNMTYGTFFNLGGRLMIPNTGISLLIP 497  
 Db 509 GTYPSDFARDTHFLHLRSASLGSQQLLGLPRDPGSSVSGTFGCLGGRLSIPGTGVSLVLP 568  
 Qy 498 PDAIPRGKIYEIYLTLLHKPEDVRLPLA-GCQTLSPIVSCGPPGVLLTRPVILAMDHCGE 556  
 Db 569 NGAIPOGKFYEMYLLINKAEST-LPLSEGTQTVLSPSVTCGPTGLLLCRPVILTMPHCAE 627  
 Qy 557 PSPDSWSLRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEAL 616  
 Db 628 VSARDWIFQLKTQAHQGHWEVVTLDEETLNTPCYQCQLEPRACHILLDQLGTYVFTGESY 687  
 Qy 617 SVAAAKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHF 676  
 Db 688 SRSVAVKRLQLAVFAPALCTSLEYSIRVYCLEDTFVALKEVLELERTLGGYLVEEPKPLMF 747  
 Qy 677 KDSYHNLRLSLHDPVSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLA 736  
 Db 748 KDSYHNLRLSLHDLPHAHWSKLLAKYQEIPFYHIWSGSQKALHCTFTLERHSLASTELT 807  
 Qy 737 CKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAG--VPALVGPSAFKIPFLIRQKI 794  
 Db 808 CKICVRQVEGEGQIFQLHTTLA-ETPAGSLDTLCSAPGSTVTTQLGPYAFKIPLSIRQKI 866  
 Qy 795 ISSLDPPCRRGADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAA 854  
 Db 867 CNSLDAENSRGNDWRMLAQKLSMDRYLNYFATKASPTGVILDLWEALQQDDGDLNSLASA 926  
 Qy 855 VAGLGQPDAGLFTVSEAE 873  
 Db 927 LEEMGKSEMLVAVATDGDC 945

# RESULT 12

## UN5D\_HUMAN

ID UN5D\_HUMAN STANDARD; PRT; 953 AA.  
 AC Q6UXZ4; Q8WYP7;  
 DT 25-OCT-2004 (Rel. 45, Created)  
 DT 25-OCT-2004 (Rel. 45, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Netrin receptor UNC5D precursor (Unc-5 homolog D) (Unc-5 homolog 4)  
 DE (UNQ6012/PRO34692).  
 GN Name=UNC5D; Synonyms=KIAA1777, UNC5H4;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,  
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,  
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,  
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,  
 RA Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D.,  
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P., Gray A.;  
 RT "The secreted protein discovery initiative (SPDI), a large-scale  
 RT effort to identify novel human secreted and transmembrane proteins: a  
 RT bioinformatics assessment.";  
 RL Genome Res. 13:2265-2270(2003).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=Brain;  
 RA Nakajima D., Nakayama M., Nagase T., Ohara O.;  
 RT "Identification of unc5H4 gene.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: Receptor for netrin. May be involved in axon guidance by  
 CC mediating axon repulsion of neuronal growth cones in the  
 CC developing nervous system upon ligand binding. Axon repulsion in  
 CC growth cones may be caused by its association with DCC that may  
 CC trigger signaling for repulsion. It also acts as a dependence  
 CC receptor required for apoptosis induction when not associated with  
 CC netrin ligand (By similarity).  
 CC -!- SUBUNIT: Interacts with the cytoplasmic part of DCC (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q6UXZ4-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q6UXZ4-2; Sequence=VSP\_011703;  
 CC Note=No experimental confirmation available;  
 CC -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By  
 CC similarity).  
 CC -!- PTM: Proteolytically cleaved by caspases during apoptosis. The  
 CC cleavage does not take place when the receptor is associated with  
 CC netrin ligand. Its cleavage by caspases is required to induce  
 CC apoptosis (By similarity).  
 CC -!- SIMILARITY: Belongs to the UNC-5 family.  
 CC -!- SIMILARITY: Contains 1 death domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC -!- SIMILARITY: Contains 1 ZU5 domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial



CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AB055056; BAB83663.1; -.  
 DR EMBL; AY358147; AAQ88514.1; -.  
 DR HSSP; P07996; 1LSL.  
 DR Genew; HGNC:18634; UNC5D.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR011029; DEATH\_like.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP\_1.  
 DR InterPro; IPR000906; ZU5.  
 DR Pfam; PF00531; Death; 1.  
 DR Pfam; PF00047; ig; 1.  
 DR Pfam; PF00090; TSP\_1; 2.  
 DR Pfam; PF00791; ZU5; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00408; IGc2; 1.  
 DR SMART; SM00209; TSP1; 2.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; FALSE\_NEG.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS50092; TSP1; 2.  
 KW Alternative splicing; Apoptosis; Developmental protein;  
 KW Immunoglobulin domain; Phosphorylation; Receptor; Repeat; Signal;  
 KW Transmembrane.

FT	SIGNAL	1	32	Potential.
FT	CHAIN	33	953	Netrin receptor UNC5D.
FT	DOMAIN	33	379	Extracellular (Potential).
FT	TRANSMEM	380	400	Potential.
FT	DOMAIN	401	953	Cytoplasmic (Potential).
FT	DOMAIN	54	151	Ig-like.
FT	DOMAIN	153	244	Ig-like C2-type.
FT	DOMAIN	252	306	TSP type-1 1.
FT	DOMAIN	308	360	TSP type-1 2.
FT	DOMAIN	540	642	ZU5.
FT	DOMAIN	859	936	Death.
FT	SITE	416	417	Cleavage (by caspase-3) (By similarity).
FT	SITE	703	721	Interaction with DCC (By similarity).
FT	DISULFID	75	134	By similarity.
FT	DISULFID	180	231	By similarity.
FT	CARBOHYD	117	117	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	228	228	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	353	353	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	376	376	N-linked (GlcNAc. . .) (Potential).
FT	VARSPLIC	1	34	MGRAAATAGGGGGGARRWLPWLGLCFWAAGTAAAR -> MIL
FT				VLVKALSDVCAGTSGFLLDFFSSQTSP (in isoform
FT				2).
FT				/FTId=VSP_011703.
SQ	SEQUENCE	953 AA;	105879 MW;	5F893B9DF746F731 CRC64;

Query Match 46.9%; Score 2185.5; DB 1; Length 953;  
 Best Local Similarity 46.3%; Pred. No. 1.2e-155;  
 Matches 421; Conservative 159; Mismatches 268; Indels 61; Gaps 12;

Qy 10 VPGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTD 69  
:| | | | | : | | : | | : | : | | | | | | | | | | : | | | : |

Db 46 IPSA-PGTLPHFIEEPDDAYIIKSNPIALRCKARPAMQIFFKCNGEWVHQNEHVSEETLD 104

Qy 70 GSSGLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQE 129  
| | | | | | | | : | | | | | | | | : | | | | | | | | : | | | | | | | | :

Db 105 ESSGLKVREVFINVTRQQVEDFHGPEDYWCQCVAWSHLGTSKSRKASVRIAYLRKNFEQD 164

Qy 130 PLAKEYSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSLDPNVYITREHSLVVRQARLA 189  
| : | | : | | | | | | | : | | | | | : | : | | : | : | | | | | :

Db 165 PQGREVPPIEGMIVLHCRPPEGVPAAEVEWLKNEEPIDSEQDENIDTRADHNLIRQARLS 224

Qy 190 DTANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSASCGRGWQKRSRSCNTPAP 249  
| : | | | | : | | | | : | | | | | | | | : | | | | | | | | : | | | | |

Db 225 DSGNYTCMAANIVAKRRSLSATVVVYVNGGWSSWTEWSACNVRCGRGWQKRSRTCTNPAP 284

Qy 250 LNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECSDPAPRNGGEEC 309  
| | | | | | | | : | | | | | | | | : | | | | | : | | | | : | | | | | :

Db 285 LNGGAFCEGMSVQKITCTSLCPVDGSWEVWSEWSVCSPECEHLRIRECTAPPPRNGGKFC 344

Qy 310 QGTDLDTRNCTSDLCV-----HSASGPEDVALYVGLIAVAVCLVLLLLVLILVY 358  
: | : : | | | | | : | | | | | | | | | | | | : | : : : : |

Db 345 EGLSQESENCTDGLCILDKKPLHEIKPQSIENASDIALYSGL-GAAVVAVAVLVIGVTLY 403

Qy 359 CRKKEGLDSDVADSSILTSGFQPVSIKPSKADNPHELL--TIQPDLSTTTTTTYQGSICPRQ 416  
| : | : | | | | | | | : | : | | | : | | | | : | | | : | | |

Db 404 RRSQSDYGVDVIDSSALTGGFQTFNFKTVRQNSLLLSAMQPD-LTVSRTYSGPIC-LQ 461

Qy 417 DGPSPKFQLTNGHLLSPLG-----GGRHTLHH 443  
| | | | : | | : | | | | | | | | | | | | | | | | | |

Db 462 D-PLDKELMTFESSLENPLSDIKVKVQSSFMVSLGVSERAEYHGKNHSRTFPHGNHNSFST 520

Qy 444 SSPTSEAEFVSRLSTQNYFRSLPRGTSNMITYGTFFNLGGRLMIPNTGISLLIPDAIPR 503  
| : : : : | | | | | | | | | | | | | | | | | | | | | |

Db 521 MHPRNKM-PYIQNLS-----SLPTRTELRTTGVEFHLGGRLVMPNTGVSLIIPHGAIP 573

Qy 504 GKIYEIYLTLLHKPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPSPDWS 563  
: | | | : : : | | | : | | | | : | | | : : | | | : | | : | : | :

Db 574 ENSWEIYMSINQGEPSLQSDGSEVLLSPEVTCGPPDMIVTTPFALTIPHCADVSEHWN 632

Qy 564 LRLKKQSCGWSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVAANKR 623  
: | | | : : | | : | : : | | | | : | | | : | | | : | | : | | :

Db 633 IHLKKRTQQGKWEVMSVEDESTS--CYCLDPFACHVLLDSFGTYALTGEPIITDCAVKQ 690

Qy 624 LKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSYHNL 683  
| | : : | : : | | : | | | : | | : | | | : | | : | | : | | : | :

Db 691 LKVAVFGCMSCNSLDYNLRVYCVDNTPCAFQEVVSDERHQGGQLLEPKLLHFKGNTFSL 750

Qy 684 RLSIHDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERVSPSTSDLACKLWVWQ 743  
: | | : | : | | : | | : | | : | | : | | : | | : | | : | | :

Db 751 QISVLDIPPFLWRIKPFTACQEVFPFSRVWCSNRQPLHCAFSLERYTPTTTQLSCKICIRQ 810

Qy 744 VEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPFLIRQKISSLDPPCR 803  
: | | : | : | : | : : | | | | | | | | : | | : | | :

Db 811 LKGHEQILQVQTSILESERETITFFAQEDSTFPAQTGPKAFKIPYSIRQRICATFDTPNA 870

Qy 804 RGADWRTLQAKLHLDLHLSFFASKPSPTAMILNLWEARHFPNGNLSQLAAAVAGLGQPD 863

```

      :| ||: ||| :: :||:|:: ||:|:||||||| :|:| || |: |:
Db      871 KGKDWQMLAQKNSINRNLSYFATQSSPSAVILNLWEARHQHGDGLDSLACALEEIGRTHT 930

Qy      864 GLFTVSEAE 872
      | :||:
Db      931 KLSNISESQ 939

```

RESULT 13

UN5D\_MOUSE

```

ID      UN5D_MOUSE      STANDARD;      PRT;      956 AA.
AC      Q8K1S2;
DT      25-OCT-2004 (Rel. 45, Created)
DT      25-OCT-2004 (Rel. 45, Last sequence update)
DT      25-OCT-2004 (Rel. 45, Last annotation update)
DE      Netrin receptor UNC5D precursor (Unc-5 homolog D) (Unc-5 homolog 4).
GN      Name=Unc5d; Synonyms=Unc5h4;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX      MEDLINE=22239710; PubMed=12351186; DOI=10.1016/S0925-4773(02)00248-4;
RA      Engelkamp D.;
RT      "Cloning of three mouse Unc5 genes and their expression patterns at
RT      mid-gestation.";
RL      Mech. Dev. 118:191-197(2002).
CC      -!- FUNCTION: Receptor for netrin involved in cell migration. May be
CC      involved in axon guidance by mediating axon repulsion of neuronal
CC      growth cones in the developing nervous system upon ligand binding.
CC      Axon repulsion in growth cones may be caused by its association
CC      with DCC that may trigger signaling for repulsion. It also acts as
CC      a dependence receptor required for apoptosis induction when not
CC      associated with netrin ligand (By similarity).
CC      -!- SUBUNIT: Interacts with the cytoplasmic part of DCC (By
CC      similarity).
CC      -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC      -!- TISSUE SPECIFICITY: Expressed in developing limb and mammary
CC      gland.
CC      -!- PTM: Phosphorylated on cytoplasmic tyrosine residues (By
CC      similarity).
CC      -!- PTM: Proteolytically cleaved by caspases during apoptosis. The
CC      cleavage does not take place when the receptor is associated with
CC      netrin ligand. Its cleavage by caspases is required to induce
CC      apoptosis (By similarity).
CC      -!- SIMILARITY: Belongs to the UNC-5 family.
CC      -!- SIMILARITY: Contains 1 death domain.
CC      -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC      -!- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC      -!- SIMILARITY: Contains 2 TSP type-1 domains.
CC      -!- SIMILARITY: Contains 1 ZU5 domain.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way

```



Db	96	VSEESLDESSGLKVREVFINVTRQQVEDFHGPEYWCQCVASHLGTSKSRKASVRIAYL	155
Qy	123	RKNFEQEPLAKEVSLEQGIVLPCRPPGIPPAEVEWLRNEDLVDPSLDPNVYITREHSIV	182
Db	156	RKNFEQDPQGREVPPIEGMIVLHCRPPEGVPAAEVEWLKNEEPIDSEQDENIDTRADHNLI	215
Qy	183	VRQARLADTANYTCVAKNIVARRRSASAAVIVVNGGWSTWTEWSVCSASCGRGWQKRSR	242
Db	216	IRQARLSDSGNYTCMAANIVAKRRSLSATVVVYVNGGWSSWTEWSACNVRCGRGWQKRSR	275
Qy	243	SCTNPAPLNGGAFCEGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWRSDPAP	302
Db	276	TCTNPAPLNGGAFCEGMSVQKITCTALCPVDGSWEVWSEWSVCSPECEHLRIRECTAPPP	335
Qy	303	RNGGEECQGTDLDRNCTSDLCV-----HSASGPEDVALYVGLIAVAVC	346
Db	336	RNGGKFCEGLSQESENCTDGLCILDKKPLHEIKPQRWSRRGIENASDIALYSGL-GAAVV	394
Qy	347	LVLILLVLILVYCRKKEGLSDVDSSILTSQFQPVSIKPSKADNPHLL--TIQPDLTSTT	404
Db	395	AVAVLVIGVTLYRRSHSDYGVDDVIDSSALTGGFQTFNFKTVRQGNLSLLNPAMQPD-LTV	453
Qy	405	TTTYQGSCLPRQDGPSPKFQLTNGHLLSPLG-----GGRH--	439
Db	454	SRTYSGPIC-LQD-PLDKELMTESSLFNPLSDIKVKVQSSFMVSLGVSEAEYHGKNHSG	511
Qy	440	-----TLHHSSPTSEAEFVSRLSTQNYFRSLPRGTSNMITYGTFFNLGGRLMIP	488
Db	512	TFPHGNRGEFSTIHPRNKT----PYIQNLS-----SLPTRTELRTTGVFHGLGGRLVMP	561
Qy	489	NTGISLLIPPDAIPRGIYEIYLTLHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVI	548
Db	562	NTGVSLIPHGAIPEENSWEIYMSINQGEF-SLQSDGSEVLLSPEVTCGPPDMLVTTFFA	620
Qy	549	LAMDHCGEPSPDSSWLSRLKKQSCEGSWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGR	608
Db	621	LTIPHCAVDSSSEHWNHLLKKRTQQGKWEVMSVEDESTS--CYCLLDPFACHVLLDSFGT	678
Qy	609	FALVGEALSVAARKRLKLLLFAPVACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLI	668
Db	679	YALTGEPITDCAVKQLKVAVFGCMSCNSLDYNLRVYCVDNTPCAFQEVISDERHQGGQLL	738
Qy	669	QEPRVLHFKDSYHNLRSLSDVPSSLWKSLLVSYQEIPFYHIWNGTQRYLHCTFTLERV	728
Db	739	EEPKLLHFKGNTFSLQVSVLDIPPFLWRIKPFACQEVVFSRVWSSNRQPLHCAFSLERY	798
Qy	729	SPSTSDLACKLWVWQVEGDGQSFSINFNITKDTRFAELLALESEAGVPALVGPSAFKIPF	788
Db	799	TPTTQLSCKICIRQLKGHEQILQVQTSILESERETITFFAQEDSTFFAQGTGPKAFKIPY	858
Qy	789	LIRQKIISLDPPCRGRADWRTLAQKLHLDHLSFFASKPSPTAMILNLWEARHFPNGNL	848
Db	859	SIRQRICATFDTNPAKGDWQMLAQKNSINRNLISYFATQSSPSAVILNLWEARHQQDGD	918
Qy	849	SQLAATAVAGLGQPDAGLFTVSEAE	872

## RESULT 14

Q7PW78

ID Q7PW78 PRELIMINARY; PRT; 876 AA.  
 AC Q7PW78;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE ENSANGP00000005212 (Fragment).  
 GN Name=ENSANGG00000004014;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AAAB01008984; EAA14755.1; -.  
 DR GO; GO:0005515; F:protein binding; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR011029; DEATH\_like.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR000906; ZU5.  
 DR Pfam; PF00531; Death; 1.  
 DR Pfam; PF00090; TSP\_1; 2.  
 DR Pfam; PF00791; ZU5; 1.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS50092; TSP1; 2.  
 FT NON\_TER 1 1  
 FT NON\_TER 876 876  
 SQ SEQUENCE 876 AA; 96301 MW; 6F9336D53E096E00 CRC64;

Query Match 22.5%; Score 1050.5; DB 2; Length 876;  
 Best Local Similarity 30.6%; Pred. No. 4.1e-70;  
 Matches 284; Conservative 138; Mismatches 368; Indels 137; Gaps 30;

QY 18 LPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKCNGEWVRQVDHVIERSTDGSSGLPTM 77  
 || |||||: |::||:| | || | || |||:| : | || ||:  
 Db 2 LPVFLVEPKGAYVMKNRPAKLYCKASHALQISFKCSGS-TKPPPTKEHHTDPHSGVQLQ 60

QY 78 EVRINVSRRQVEKVFGLLEEWCVQVWSSSGTTKSQKAYIRIARLRKNFEQEPLAKEVSL 137  
 | ::||: |:: || : |::| |::| |::| |::| |:  
 Db 61 EATATITRELVDFFGKGPFKCECRAYSSRGHVKTQPVTIQVATIKKQISISP KIVRVAT 120

QY 138 EQGIVLPCRPPPEGIPPAEVEWLRNEDLVPSLDENVYITREHSLVVRQARLADTANYTCV 197  
 | | | |::| ||:| | : | : |::|:: : : | |||||  
 Db 121 GGRAELNC-IANATPAKVWLKNS--VPVHANPPFVLLTENALLIARVEIQDMANYTCV 177

Qy	198	AKNIVARRRSASAIVYVNGGWSTWTEWSVCSA-SCGRGWQKRSRCTNPAPLNGGAF	256
Db	178	AENIAGKRVS	237
Qy	257	EGQNVQKTACATLCPVDGWSWPWSKWSACGLDCTHWSRECS	313
Db	238	KGASTESTPDCLPCSA-GRWSSWSEWSECGPDCTQIRQRSCVAQAF	296
Qy	314	LDTRNCTSDLCVHSASGPE-DVALYVGLIAVAVCLVLLLVILVYCRKKEGLDSDVADS	372
Db	297	QOSIKCTGGLCNYTAQDSNWSVYLWVTLVA-AFCLGVVFAV--SKFLRRKKTIPAYNLAR	353
Qy	373	SILTSGFQPVSIK-PSKAD-NPHLLTIQP-----DLSTTTT	407
Db	354	SDLTQNTGPIYNEYPMASLQPHLAGHHPHHHHLQQQQHHHHLSSLHHQHGSLLGPVAT	413
Qy	408	YQGS---CPRQDGPSPKFQLTNGHLLSPLGGGRHTL-----HHSS-----	445
Db	414	GHGQLHPQCQSQQAPT-----LPIGGLKSSLPLPRSNSEHHYDVPHL	462
Qy	446	--PTSEAEFVSRLSTQNYFRSLPRGTSNMITYGTNFGGLRMI	503
Db	463	IRPTTLVEE-----PFRGAEVTHATLT	508
Qy	504	GKIYEIYLT	561
Db	509	HQRHVSVALSIVRDDKHHPVPPTGPRSTYLS	567
Qy	562	WSLRLKKQSCEGS-WEDVLHLGEEAPSHLYYCOLEASAC	617
Db	568	WAFSLYSAPDNVTPWCKVVTIGEETLNTPALVQIDKRYAYVLTETFGKYVLVGESATDIQ	627
Qy	618	-VAAAKRLKLLL	675
Db	628	ERVACKRLRLFICGPSTVPEFSDVSLRVYIVEDNPGAEE	687
Qy	676	FKD-SYHNLRSLIHDPSSLWKSLLVSYQEI	734
Db	688	FADVGOAGLNIDLQCVGG--WRAKSSSERQTIPFSHVWNSACTALHCSFTL-----	736
Qy	735	LACKLWVWQVEGDGQSFSINFNITKDTRFA---ELLALESEAGVPA-----	777
Db	737	--CR-----TEHDKCDFKIVVQASQDVPQGLDERLTAI----	785
Qy	778	-LVGPSAFKIPFLIRQKIISLDPFCRRGADWRTLAQKLHLD	836
Db	786	NLVATDRFRLSKDVKRKLCRCLDPPTQKRNDWRMLAAHLNVDRYLTYFATRPSPTDQILD	845
Qy	837	LWEARHFPNGNLSQLAAAVAGLGQ	863
Db	846	LWEARNRDLNALOOLIEICRTMER	872

ID UNC5 DROME STANDARD; PRT; 1072 AA.

AC Q95TU8; Q9NBL0; Q9V7B5;  
DT 25-OCT-2004 (Rel. 45, Created)  
DT 25-OCT-2004 (Rel. 45, Last sequence update)  
DT 25-JAN-2005 (Rel. 46, Last annotation update)  
DE Netrin receptor unc-5 precursor (Unc5 netrin receptor).  
GN Name=unc-5; ORFNames=CG8166;  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.  
RX PubMed=11719202;  
RA Keleman K., Dickson B.J.;  
RT "Short- and long-range repulsion by the Drosophila Unc5 netrin  
RT receptor.";  
RL Neuron 32:605-617(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,



RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*."; .  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP GENOME REANNOTATION.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review."; .  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley; TISSUE=Embryo;  
 RX MEDLINE=22426066; PubMed=12537569;  
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
 RA George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,  
 RA Rubin G.M., Celniker S.E.;  
 RT "A *Drosophila* full-length cDNA resource."; .  
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
 RN [5]  
 RP FUNCTION.  
 RX PubMed=12765609;  
 RA Freeman M.R., Delrow J., Kim J., Johnson E., Doe C.Q.;  
 RT "Unwrapping glial biology: Gcm target genes regulating glial  
 RT development, diversification, and function."; .  
 RL Neuron 38:567-580(2003).  
 CC -!- FUNCTION: Receptor for netrin required for motor axon guidance.  
 CC Mediates both short- and long-range axon motor repulsion in the  
 CC developing nervous system upon ligand binding. Also involved in  
 CC glial migration. While short-range repulsion requires both fra and  
 CC unc-5, long-range repulsion only requires unc-5.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -!- TISSUE SPECIFICITY: Prior to gastrulation, it is strongly  
 CC expressed in the presumptive mesoderm. Mesodermal expression  
 CC begins to fade during stages 13-14, persisting only in the cells  
 CC that form the dorsal vessel. Expressed within the CNS from late  
 CC stage 13, shortly after the first axons have extended. Detected in  
 CC several dispersed clusters of cells within the CNS, increasing in  
 CC number as development proceeds. Also expressed in the peripheral  
 CC and exit glia, which migrate laterally out of the CNS between  
 CC stages 14 and 17. Strongly expressed in motor axons that exit the  
 CC CNS ipsilaterally via the segmental nerve root (SN). Not expressed  
 CC on either commissural or longitudinal axons within the CNS, nor on  
 CC motor axons that exit via the intersegmental nerve (ISN). In the  
 CC periphery, it is detected on all branches of the SN. Also  
 CC expressed at high level in exit and peripheral glia along both the  
 CC SN and ISN.  
 CC -!- PTM: Phosphorylated on different cytoplasmic tyrosine residues (By  
 CC similarity).  
 CC -!- SIMILARITY: Belongs to the UNC-5 family.  
 CC -!- SIMILARITY: Contains 1 death domain.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.  
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like domain.  
 CC -!- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF247762; AAF74193.1; -.  
 DR EMBL; AE003811; AAF58143.2; -.  
 DR EMBL; AY058501; AAL13730.1; -.  
 DR HSSP; P07996; 1LSL.  
 DR FlyBase; FBgn0034013; unc-5.  
 DR GO; GO:0007411; P:axon guidance; IGI.  
 DR GO; GO:0008347; P:glia cell migration; IMP.  
 DR InterPro; IPR000488; Death.  
 DR InterPro; IPR011029; DEATH\_like.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR InterPro; IPR000884; TSP1.  
 DR Pfam; PF00531; Death; 1.  
 DR Pfam; PF00047; ig; 1.  
 DR SMART; SM00005; DEATH; 1.  
 DR SMART; SM00408; IGc2; 1.  
 DR SMART; SM00209; TSP1; 2.  
 DR PROSITE; PS50017; DEATH\_DOMAIN; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 DR PROSITE; PS50092; TSP1; 2.  
 KW Developmental protein; Immunoglobulin domain; Phosphorylation;  
 KW Receptor; Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 30 Potential.  
 FT CHAIN 31 1072 Netrin receptor unc-5.  
 FT DOMAIN 31 440 Extracellular (Potential).  
 FT TRANSMEM 441 461 Potential.  
 FT DOMAIN 462 1072 Cytoplasmic (Potential).  
 FT DOMAIN 128 224 Ig-like.  
 FT DOMAIN 232 314 Ig-like C2-type.  
 FT DOMAIN 324 379 TSP type-1 1.  
 FT DOMAIN 398 499 TSP type-1 2.  
 FT DOMAIN 980 1067 Death.  
 FT DISULFID 149 207 By similarity.  
 FT DISULFID 253 303 By similarity.  
 FT CARBOHYD 79 79 N-linked (GlcNAc . .) (Potential).  
 FT CARBOHYD 300 300 N-linked (GlcNAc . .) (Potential).  
 FT CONFLICT 885 885 P -> S (in Ref. 1).  
 SQ SEQUENCE 1072 AA; 116416 MW; A0A6B5A96B10138F CRC64;

Query Match 21.3%; Score 992; DB 1; Length 1072;  
 Best Local Similarity 28.7%; Pred. No. 1.4e-65;  
 Matches 282; Conservative 145; Mismatches 383; Indels 174; Gaps 33;

Qy 11 PGANPDLLPHFLVEPEDVYIVKNKPVLLVCKAVPATQIFFKNGEWVRQVDHVIERSTDG 70  
 || : || ||:|||| |:||||:| :| ||| : |: |||:| : | |

Db 120 PGEASNTLPIFLIEPESVFWVKNRPAVLKCKASHSLQVIFKCSGS-SQPPPSTHETHVDP 178  
 Qy 71 SSSLPTMEVRINVSRRQVEKVFGLLEEYWCQCVAWSSSGTTKSQKAYIRIARLRKNFEQEP 130  
 :|: || : | |:: || : |:: |||| | ||| | : || :||:| | |  
 Db 179 HTGVNMEEVTATIHRDLVDEFFGDGPFKCECHAWSSRGVVKSAATVHIAYIRKSFNQSP 238  
 Qy 131 LAKEVSLEQGIVLPCRPEGIPPAEVEWLRNEDLVDPSPDPNVYITREHSLVVRQARLAD 190  
 : : | | | | | : | : | : : : | : | | |  
 Db 239 TSLRLELGSRAELRCEPPGGFPEPKLTWHKNNAVITADSEPGITVS-AGTLIFRQVALQH 297  
 Qy 191 TANYTCVAKNIVARRRSASAAVIVYVNGGWSTWTEWSVCSA----SCGRGWQKRSRSCN 246  
 |||:| |::| || | | :||| |||||: | | | || | |||:| |  
 Db 298 MANYSCSAENIAGRRVSDSAVLIVYVNGGWSTWSPWRECKCAGKPSQGR---KRSRTCNN 354  
 Qy 247 PAPLNGGAFCEGQNVQKTACATLCPVD-----GSWPWSKWSACGLDC 289  
 | ||||| | | :||:| || | || || | | :|  
 Db 355 PMPLNGGAQCPCGPQIQKSADCAACPEDTQIVSPDGFDISSSKRMARWSAWSWSICSAC 414  
 Qy 290 THWRSREC-----SDP 300  
 | |::| ||  
 Db 415 IQVRRRKCLTQGQTQISSEAEAEAGDLLLGAPGVGMAALIAAGVGAVGSPSEATGSSSDI 474  
 Qy 301 APRNGGEECQGTDLDRNCTSDLCVHSASGPEDV--ALYVGLIAV-AVCLVLLLLVLILV 357  
 | | | | | : | | : | | : | | : || : ||| : || :  
 Db 475 IPGYGKSLCAGKDIQTAECRGEQC---QIGKDDFDWTLYLGLAFITAVCFAGTALI--- 528  
 Qy 358 YCRKKEGLDSDV---ADSSILTSGFQPVSIKPSKADNPHLLTIQPDLSSTTTTTYQGS LCP 414  
 | : | : : : | : : : | : | : | : |  
 Db 529 --CCARRGIRTNPHYNMARSVMDADYMP-GVVEKKEMRMHI-----EASNMGYDYV----- 576  
 Qy 415 RQDGPSPKFQLTNGHLLS-PLGGGRHTLHH-----SSPTSEAE 452  
 | : : | | : : : | | | | | | | : | :  
 Db 577 ---NPGHRY-LPGEHIMGMGIGCGGVTEHHYDVPNLSANYTNPIDHLSVDYLSETGESST 632  
 Qy 453 FVSRLST--QNYFRSLPRGTSNMTYGTFNFLGGRLMIPNTGISLLIPDAIPRGKIYEIY 510  
 : || | | : : : || ||:| : : | : | | : :  
 Db 633 ADTSNSTFDMNGKLSILNASKSSTYEMLGSAAGQLRLYGGELLFVPEHAIGKHVKKHVS 692  
 Qy 511 LTLHKPEDVRLPLAGCQT----LLSPIVSCGPPGVLLTRPVILAMDHCGEPSDWSLRL 566  
 | | | | : | | | : | | : ||| : || : | : :  
 Db 693 LLLLSECSRV---SCATESSILCSSVVSAPRNYSFVKPVILKIPHC-LVAPEQWHVHI 748  
 Qy 567 KKQSCEG-----SWEDVLHLGEEAPSHLYYCQLEASACYVFTEQLGRFALVGEALSVA 621  
 | : | : : ||| : : ||||: : |||| | : | :  
 Db 749 YHADSEHDELSVNWRRAVSVEETINTPMFVQLEATHVFIMTEQLGHFTVVAEPRIQQPS 808  
 Qy 622 KRLKLLLFAP-VACTSLEYNIRVYCLHDTHDALKEVVQLEKQLGGQLIQEPRVLHFKDSY 680  
 :||| | : : : :| : | : : : | : || : | : | :  
 Db 809 IKMKLLAFSQHTPPSNANCSLRIYVVKDFPNRDI CANVEAKLGGFSGESQVFAFTLNS 868  
 Qy 681 HNLRLSIH--DVPSSLWKSLLVSYQE-IPFYHIWNGTQRYLHCTFTLERVSPSTSDIAC 737  
 || : : || : : | : ||: || : || | : | : : |  
 Db 869 RNLNIRVRSADVEAA-----APYEHAIPYQHILSNNS-ILHCEFSLRQDQNS---LC 917  
 Qy 738 KLWWVQVEGDGQSFSINFNITKDTRFAELLALESEAGVPA-----LVGPSA----FKIPF 788  
 : | | :|| : : |||| : : | | : :|  
 Db 918 VDFGQGSSEDDYYTFNIPAHSMMSG--AEELASTTNTTISIDRQGNVNESCVMDFVQLPH 975

Qy 789 LIRQKIISLDPPCRRGADWRTLAQKLHLDSHLSFFASKPSPTAMILNLWEAR--HFPNG 846  
 :: | :||| ||| ||:||: | ::::||:| ||| ||||| | |  
 Db 976 ATKRLICGALDPPRADERDWRLAKKLNTDRYIAYFATKASPTEQILNLWECRANSSPGS 1035  
 Qy 847 NLSQLAAAVAGLGQPDAGLFTVSE 870  
 : : :: : | | : |  
 Db 1036 SSNSVSHTIMAL-----LLTLKE 1053

Search completed: March 1, 2005, 09:03:33  
 Job time : 162.7 secs